8-11-2021

Waveform systematics in parameter inference of GW signals from compact binary mergers

Anjali Balasaheb Yelikar

RIT

Waveform systematics in parameter inference of GW signals from compact binary mergers

By

Anjali Balasaheb Yelikar

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science. in Astrophysical Sciences & Technology

School of Physics and Astronomy
College of Science
Rochester Institute of Technology
Rochester, NY

Aug 11, 2021

Approved by:_____

Andrew Robinson, Ph.D Date
Director, Astrophysical Sciences and Technology

ASTROPHYSICAL SCIENCES AND TECHNOLOGY
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

MASTER DEGREE THESIS

The Master's. Degree Thesis of Anjali Yelikar has been examined and approved by the thesis committee as satisfactory for the dissertation requirement for the Master of Science degree in Astrophysical Sciences and Technology.

Dr. Richard O'Shaughnessy, Thesis Advisor

Dr. Joshua Faber, Committee Member

Dr. Jason Nordhaus, Committee Member

Date _____

Declaration

I, ANJALI BALASAHEB YELIKAR ("the Author"), declare that no part of this thesis is substantially the same as any that has been submitted for a degree or diploma at the Rochester Institute of Technology or any other University. I further declare that the work in Chapter 3 is entirely my own; Chapter 2 is drawn from work done with group members at Rochester Institute of Technology; Chapters 1 and 4 draw in parts on work done as a part of the LIGO Scientific Collaboration, including other authors contributions. Those who have contributed scientific or other collaborative insights are fully credited in this thesis, and all prior work upon which this thesis builds is cited appropriately throughout the text. I have omitted reference to my contributions to other papers by the LIGO Scientific, Virgo, KAGRA Collaboration, such as GWTC-2.1 catalog [1].

- Chapter 2 is based on and in parts directly drawn from the paper [2], entitled *Assessing and marginalizing over compact binary coalescence waveform systematics with RIFT* (2020,PRD,102(12):124069) authored by A. Z. Jan, A. B. Yelikar, J. Lange, and R. O'Shaughnessy., arXiv:2011.03571

- Chapter 4: Section 4.1 is based on and in parts directly drawn from the paper [3], entitled *Observation of Gravitational Waves from Two Neutron Star–Black Hole Coalescences* (2021,Astrophys. J. Lett.,915(1):L5) authored by the LIGO Scientific, Virgo, KAGRA Collaboration, arXiv:2106.15163

- Chapter 4: Section 4.2 is based on and in parts directly drawn from the paper [4], entitled *Search for intermediate mass black hole binaries in the third observing run of Advanced LIGO and Advanced Virgo* (2021:LIGO-P2100025) authored by the LIGO Scientific, Virgo, KAGRA Collaboration: arXiv:2105.15120

Acknowledgements

I thank my advisor Dr. Richard O'Shaughnessy for his guidance and support throughout the project, my collaborators Aasim Z. Jan and Dr. Jacob Lange for their discussions and inputs. I would also like to thank my Dr. Joshua Faber and Dr. Jason Nordhaus for agreeing to be on my MS committee. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. I am grateful to the computing resources provided by California Institute of Technology, LIGO Hanford Observatory, LIGO Livingston Observatory, and the Inter-University Centre for Astronomy and Astrophysics, India.

I am thankful for the support of my family and friends, in this strange time we all are living through.

Abstract

The Nobel prize winning discovery of gravitational waves from a binary black hole merger GW150914 opened up a new window onto the Universe. We have now seen multiple GW detections from coalescences of different kinds of compact binary objects. Accurate inference of parameters of these compact objects is a crucial part of gravitational wave astronomy. Data analysis techniques employ Bayesian statistics comparing gravitational wave models against the detected signal. Most of these models approximate solutions of Einstein's General Relativity equations, as generating numerical relativity(NR) solutions for every point in the parameter space of probable compact binary coalescences is computationally expensive. The equations are hence generally solved using analytical or semi-analytical approximations and then compared to existing NR simulations in the most nonlinear and dynamical regime. These models are subject to waveform modeling uncertainties or systematics. In this work, we provide example(s) of these systematic differences pertaining to gravitational waveform models describing mergers of compact objects and propose an efficient technique to marginalize over these differences for a given set of waveform models. We also investigate systematic differences between tidal waveform models that include higher-order modes, quantifying the differences between the inclusion and omission of higher-order modes. The marginalization technique in combination with our very efficient parameter inference algorithm RIFT, can directly account for any available models, including very accurate but computationally costly waveforms. I also describe several contributions to results performed as a part of the LIGO Scientific Collaboration, including the interpretation of the first discovered BHNS binaries.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Gravitational waves

Gravitational waves (GW) were predicted by Albert Einstein in 1916 in his General Theory of Relativity (GR). They are propagating oscillations of the gravitational field and are generated by accelerating masses just as electromagnetic waves are generated by accelerating electrically-charged particles. They interact very weakly with matter as gravity is the weakest of the four fundamental forces unlike electromagnetic waves which interact with the medium they pass through. In simple terms, a weak gravitational field can be defined as one in which spacetime is nearly flat (Minkowskian), and the gravitational waves being propagating perturbations in this field can be expressed as:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{1}$$

where $\eta_{\mu\nu}$ = diag(-1,1,1,1) is the flat metric and $h_{\alpha\beta} \ll 1$.

The Einstein field equations relate the local spacetime curvature (Einstein tensor-$G_{\mu\nu}$) to the local energy-momentum stress tensor ($T_{\mu\nu}$). Inserting $g_{\mu\nu}$ in Einstein's tensor we get

$$\frac{8\pi G T_{\mu\nu}}{c^4} = G_{\mu\nu} = \frac{1}{2}(\partial_\sigma \partial_\mu h^\sigma_\nu + \partial_\sigma \partial_\nu h^\sigma_\mu - \partial_\mu \partial_\nu h - \Box h_{\mu\nu} - \eta_{\mu\nu}\partial_\rho \partial_\lambda h^{\rho\lambda} + \eta_{\mu\nu}\Box h) \tag{2}$$

Using Lorenz gauge ($\partial_\mu h^\mu_\nu$) and the "trace-reversed" metric ($\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}$), the expression simplifies to

$$\Box \bar{h}_{\mu\nu} = \frac{-16\pi G T_{\mu\nu}}{c^4} \tag{3}$$

The above equation can be solved using wave solutions. When a GW passes a ring of test masses it sets them into motion transverse to the direction of propagation. The gravitational wave amplitude/strain h can be decomposed using basis tensors (e_+ and $e_\times$) in free space as

$$h_{\mu\nu} = h_+ e_{+,\mu\nu} + h_\times e_{\times,\mu\nu} \tag{4}$$

These basis tensors describe the two modes of polarizations of GW- plus(e_+) and cross($e_\times$) polarization. The gravitational wave in Fig. 1 has plus polarization only.

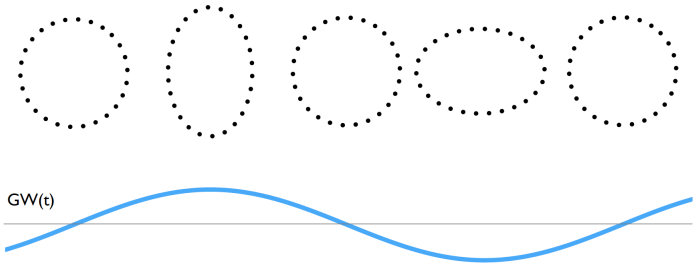

Figure 1: Effect of GWs passing transverse to the plane of the paper, on a ring of particles.

1.2 Gravitational wave detectors

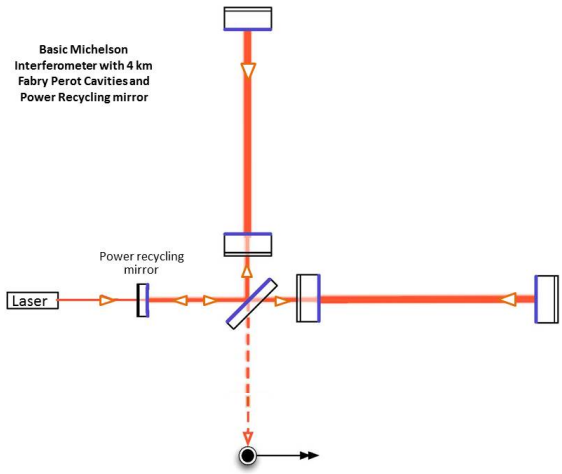

Figure 2: Working of a ground-based GW detector.

The Laser Interferometer Gravitational wave Observatory (LIGO) and Virgo experiments are the ground-based detectors searching for gravitational waves in the frequency range 10-2000Hz. There are two LIGO detectors in USA and one Virgo detector in Italy. A fourth detector KAGRA, located underground in Japan has recently joined this international effort and will help in increasing and improving our sky coverage and

localization.

The basic working principle of these detectors is that of a Michelson interferometer where the LIGO detectors have an individual arm length of 4km. A high-power laser is split into the two arms using a beam-splitter and after multiple reflections from the mirrors at the ends of the arms, they interfere at the beam-splitter and enter a photodiode. Multiple reflections are required as the strain (h) of the incoming GW is very small ($h = \frac{\Delta L}{L}$), where L is the effective path length the laser light has travelled, and for the detector to be sensitive to this change, the effective path length the light travels needs to be increased. If no GW wave passes through the interferometer, the path lengths in the two arms are the same and the interferometer is setup to get destructive interference at the detector. If a GW passes through it, the path lengths will be different for that fraction of time and the photodiode records this light in the form of voltage.

As gravitational waves couple very weakly to our detectors, those astronomical sources that we can detect must be extremely luminous in gravitational radiation. Gravitational wave astronomy therefore is biased toward looking for highly energetic, even catastrophic, events.

Gravitational waves were first detected from a merger of two black holes on September 14, 2015 by the two Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) detectors at Hanford and Livingston in USA [5]. The event GW150914, was a result of a $35.6M_{\odot}$ and a $30.6M_{\odot}$ black hole (BH) inspiralling and merging to form a $63.1M_{\odot}$ black hole and radiating away the remaining mass in the form of gravitational waves. The first merger of two neutron stars was observed on 17 Aug, 2017-GW170817 [6] and a subsequent one, GW190425 [7] was detected about two years later. In January of 2020, two confident detections of neutron star-black hole mergers were made- GW200105 and GW200115 [3]. Since then **50** gravitational wave events have been reported by the LIGO-Virgo-Kagra (LVK) collaboration from the first two observing runs (O1,O2) [8] and first-half of the third observing run (O3a) [9], albeit with no electromagnetic counterpart detection reported for either. An update to this catalog of events was released recently [1], which makes use of improved calibration, better subtraction of excess noise and an improved search criterion, taking the total up to **55** events.

1.3 Gravitational wave sources

Gravitational waves are emitted by sources that have quadrupole moment (Q) or higher moments and are dynamical. The leading order is quadrupolar, as described by the quadrupole formula: $h \sim \frac{1}{r}\frac{d^2Q}{dt^2}$, where h is the GW amplitude/strain. These two conditions ensure the second derivative of quadrupole moment to be non-zero.

Sources such as accreting neutron stars (NS) or excitation of various modes of a neutron stars, etc emit GW continuously with no frequency change likely during the period of observations, aptly named continuous/monochromatic gravitational waves. Supernova explosions or late-stages of compact binary mergers result in gravitational waves that are only a few milliseconds-seconds in duration. They occur as a result of core-collapse, and can help constrain equation of state (EOS) and rotation of the core. In these kind of sources we have the unique possibility to detect electromagnetic, gravitational wave and neutrino emission. Similar to the Cosmic Microwave Background (CMB) in the electromagnetic spectrum, gravitational waves can also originate from a stochastic GW background that could be of primordial or astrophysical origin. Primordial background arises from some of the fundamental processes which occurred in the very early phase of the universe soon after the Big Bang: e.g. Inflation. It is a unique way to probe the Universe when it was very young, $\sim 10^{-30}$ seconds old. Astrophysical stochastic background contribution is due to the enormous number of sources whose signals reach the detectors, but they are so frequent and signal to noise ratio is low that detector cannot resolve the individual signals. Lastly, mergers of compact object binaries such as BH-BH, NS-NS or NS-BH are also sources of gravitational waves, lasting milliseconds to seconds in duration depending on the mass of the binary. These compact binary coalescences (CBC) are detectable by ground-based GW interferometers such as LIGO,Virgo and KAGRA.

Binary black-hole (BBH) systems are described by 15 parameters: 8 intrinsic and 7 extrinsic. The intrinsic parameters uniquely define the binary's dynamics: $\vec{\lambda} = (\mathcal{M}, q, \chi_{1x}, \chi_{1y}, \chi_{1z}, \chi_{2x}, \chi_{2y}, \chi_{2z})$, and extrinsic parameters characterize its spacetime location and orientation: $\vec{\theta} = (t, \phi, d_L, \iota, \psi, \alpha, \delta)$, where

- $\mathcal{M}$ is chirp mass of the system expressed in solar mass units, defined as $\frac{(m_1 m_2)^{3/5}}{(m_1+m_2)^{1/5}}$, m_1 being the mass of the primary (heavier) object and m_2 being the mass of the secondary (lighter) object
- q is the mass-ratio of the system, defined as $q = \frac{m_2}{m_1}$ and its value is between (0,1]

- $\chi_{1x}, \chi_{1y}, \chi_{1z}$ are the dimensionless spin-parameters of the primary object in the respective three dimensional axes
- $\chi_{2x}, \chi_{2y}, \chi_{2z}$ are the dimensionless spin-parameters of the secondary object. The individual values are in between [-1,1]
- t is the coalescence time
- ϕ is the phase at coalescence instant
- d_L is the luminosity distance of the binary source
- ι is the inclination angle of the source, the angle between the orbital angular momentum of the binary and the line-of-sight
- ψ, α, δ are the Euler angles describing the binary's orientation.

In a binary neutron star (BNS) system, the tidal field of the companion induces a mass-quadrupole moment. The ratio of the induced quadrupole moment to the external tidal field is proportional to the tidal deformability (or polarizability), $\lambda = (2/3)k_2[(c^2/G)(R/m)]^5$, where k_2 is the second Love number and R is the stellar radius.R and k_2 are fixed for a given stellar mass m by the equation of state (EOS) for neutron-star matter [6]. Hence, we have two additional intrinsic parameters, called tidal deformability parameters (λ_1, λ_2) to describe the effects of tidal gravitational field, one for each object, increasing the number of parameters describing a BNS system to 17. A widely used spin-parameter, χ_{eff}, is a dimensionless combination of masses and spins as follows $\chi_{eff} = \frac{m_1\chi_{1z}+m_2\chi_{2z}}{m_1+m_2}$ and is better constrained than the individual spins of the objects, also constrained between [-1,1].

1.4 Statistics

Gravitational wave signals are very weak and hence require optimized statistical methods of signal extraction to identify signals convoluted with detector noise. The detector or instrumental noise is considered stationary noise, i.e., its statistical properties do not change with time and is modeled as a Gaussian random process. When the statistical properties of the noise process are known and the exact form of the signal is also known,

it is possible to construct an optimal detection statistic, which quantifies the probability of a GW signal being present in the data [10].

A GW strain can be mathematically represented as

$$s(t) = h(t) + n(t) \tag{5}$$

where $s(t)$ is the signal output from the detector once a gravitational wave(s) passes through it, $n(t)$ is the detector noise realization present in the signal, $h(t)$ is the hypothesis/model which we assume explains the GW signal. We also define the noise-weighted inner-product (a, b) of two time-series $a(t)$ and $b(t)$ as

$$(a, b) = 4Re \int_0^\infty \frac{\tilde{a}(f)\tilde{b^*}(f)df}{S_n(f)} \tag{6}$$

where $S_n(f)$ is the one-sided power spectral density of the noise, a quantity similar to what is plotted in Fig. 3. We shall assume two hypotheses that describe the signal to be a) *null hypothesis* $\mathcal{H}_0 : s(t) = n(t)$ and b) *alternate hypothesis* $\mathcal{H}_1 : s(t) = n(t) + h(t)$. Null hypothesis says that the signal consists only of a random noise process, whereas the alternate hypothesis also includes a known form of GW signal. We distinguish between the two hypotheses by computing the odds ratio $O(\mathcal{H}_1|s) = P(\mathcal{H}_1|s) : P(\mathcal{H}_0|s)$ i.e., the ratio of the probability that the alternative hypothesis $\mathcal{H}_1$ is true, given the data $s(t)$, to the probability that the null hypothesis $\mathcal{H}_0$ is true given the data. To compute the odds ratio we use Bayes' theorem, A and B being two hypotheses:

$$P(B|A) = \frac{P(B) \times P(A|B)}{P(A)} \tag{7}$$

- $P(B)$ is the prior probability or marginal probability of B being true.
- $P(A)$ is the marginal probability of A , which is known as the evidence and acts as a normalizing constant.
- $P(B|A)$ is the posterior probability of B being true given that A is true or conditional probability.

Now, expressing Bayes' theorem in a more convenient form by employing the completeness relation $P(A) =$

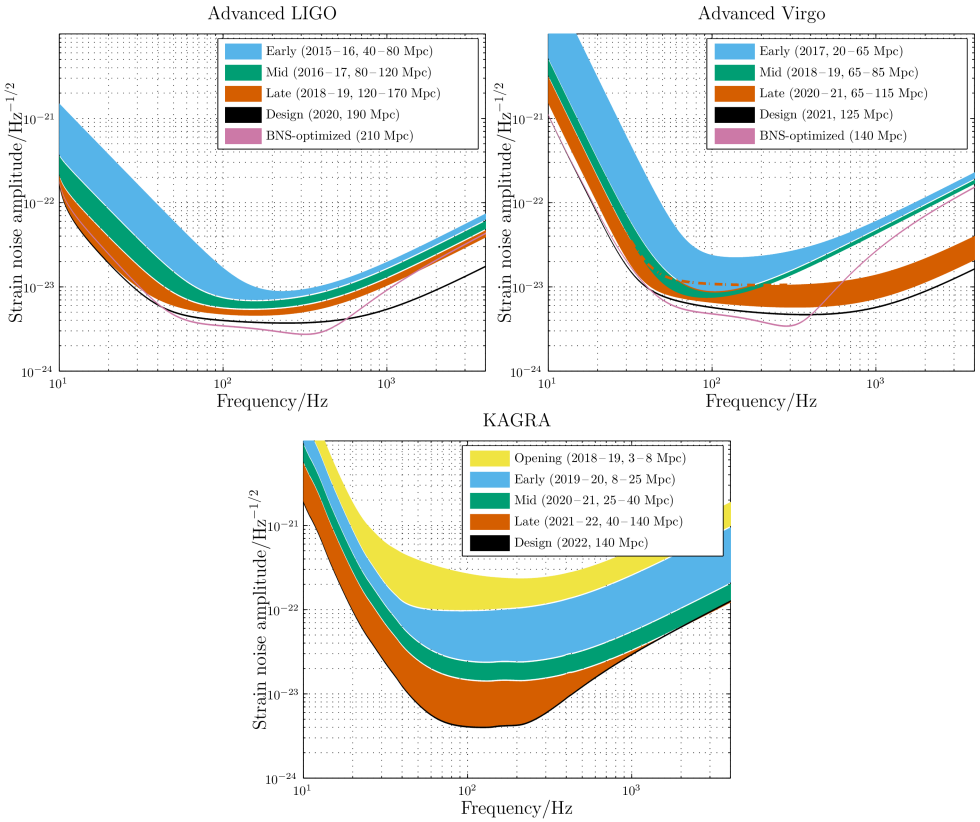

Figure 3: Sensitivity of current GW detectors. The lower the curve, the more sensitive we are in measuring gravitational waves in those frequencies.

$P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$ where $P(A|\bar{B})$ is the probability of A given that B is not true and $P(\bar{B}) = 1 - P(B)$, we get

$$P(B|A) = \frac{P(B)P(A|B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} = \frac{\Lambda(B|A)}{\Lambda(B|A) + P(\bar{B})P(B)} \tag{8}$$

$\Lambda(B|A) := \frac{P(A|B)}{P(A|\bar{B})}$ is called the likelihood ratio. The odds ratio mentioned from before can hence be expressed as

$$O(B|A) = O(B)\Lambda(B|A) \tag{9}$$

For the problem of detection we wish to decide between the two hypotheses of $\mathcal{H}_0$ and $\mathcal{H}_1$. We do this by computing the odds ratio for the alternative hypothesis given the observed data $s(t) : O(\mathcal{H}_1|s)$. Using Eq. 9

we get,

$$\Lambda(\mathcal{H}_1|s) = \frac{p(s|\mathcal{H}_1)}{p(s|\mathcal{H}_0)} \tag{10}$$

If the noise is Gaussian then the probability densities can be computed under the null hypothesis $\mathcal{H}_0$: $n(t) = s(t)$ as

$$p(s|\mathcal{H}_0) = p_n[s(t)] \propto e^{-(s,s)/2} \tag{11}$$

Under the alternative hypothesis, $\mathcal{H}_1$: $n(t) = s(t) - h(t)$ it can be expressed as

$$p(s|\mathcal{H}_1) = p_n[s(t) - h(t)] \propto e^{-(s-h,s-h)/2} \tag{12}$$

Therefore,

$$\Lambda(\mathcal{H}_1|s) = \frac{e^{-(s-h,s-h)/2}}{e^{-(s,s)/2}} = e^{(s,h)}e^{-(h,h)/2} \tag{13}$$

From above, we see that the likelihood ratio $\Lambda(\mathcal{H}_1|s)$ depends on the data $s(t)$ only through the inner product (s, h) and it is a monotonically increasing function of this inner product, hence so is the odds ratio $O(\mathcal{H}_1|s)$. Hence the optimal detection statistic is called the **matched filter** and is expressed as the following inner product:

$$(s, h) = 4Re \int_0^\infty \frac{\tilde{s}(f)\tilde{h^*}(f)df}{S_n(f)} \tag{14}$$

The matched-filter technique is the main procedure applied in model-dependent GW search algorithms such as PyCBC [11] and GstLAL [12]. There exists another GW search technique called cWB [13] that uses minimal assumptions on signal morphology to detect and reconstruct GW transients. It identifies co-incident energy across the network of detectors to classify GW signals. Once a rough idea of the signal parameters are obtained by Search methods, parameter estimation is performed to obtain accurate values of the source parameters. In this work, we utilize one such tool- Rapid parameter inference via Iterative FiTting (RIFT) [14, 15].

1.4.1 Rapid parameter inference via Iterative FiTting (RIFT)

A coalescing compact binary in a quasi-circular orbit can be completely characterized by its intrinsic and extrinsic parameters, denoted by λ and θ respectively.

RIFT [14] consists of a two-stage iterative process to interpret gravitational wave data s via comparison to predicted gravitational wave signals $h(\lambda, \theta)$. In one stage, for each λ_α from some proposed "grid" $\alpha = 1, 2, \ldots N$ of candidate parameters, RIFT computes a marginal likelihood

$$\mathcal{L}_{\rm marg} \equiv \int \mathcal{L}(\boldsymbol{\lambda}, \theta) p(\theta) d\theta \tag{15}$$

from the likelihood $\mathcal{L}(\boldsymbol{\lambda}, \theta)$ of the gravitational wave signal in the multi-detector network, accounting for detector response; see the RIFT paper for a more detailed specification. In the second stage, RIFT performs two tasks. First, it generates an approximation to $\mathcal{L}(\lambda)$ based on its accumulated archived knowledge of marginal likelihood evaluations $(\lambda_\alpha, \mathcal{L}_\alpha)$. This approximation can be generated by gaussian processes, random forests, or other suitable approximation techniques. Second, using this approximation, it generates the (detector-frame) posterior distribution

$$p_{\rm post} = \frac{\mathcal{L}_{\rm marg}(\boldsymbol{\lambda}) p(\boldsymbol{\lambda})}{\int d\boldsymbol{\lambda} \mathcal{L}_{\rm marg}(\boldsymbol{\lambda}) p(\boldsymbol{\lambda})}. \tag{16}$$

where prior $p(\boldsymbol{\lambda})$ is the prior on intrinsic parameters like mass and spin. The posterior is produced by performing a Monte Carlo integral: the evaluation points and weights in that integral are weighted posterior samples, which are fairly resampled to generate conventional independent, identically-distributed "posterior samples." For further details on RIFT's technical underpinnings and performance, see [14, 16, 15].

1.5 Gravitational waveform models

Accurate inference of source parameters is crucial to the field of gravitational wave astronomy as they are further used in studies such as tests of General Relativity (TGR) [17, 18, 19, 20], constraining equation of state (EOS) of the neutron star, forming models that describe the populations of compact object mergers [21, 22] and their formation channels [23, 24], and measuring the Hubble's constant using GW standard

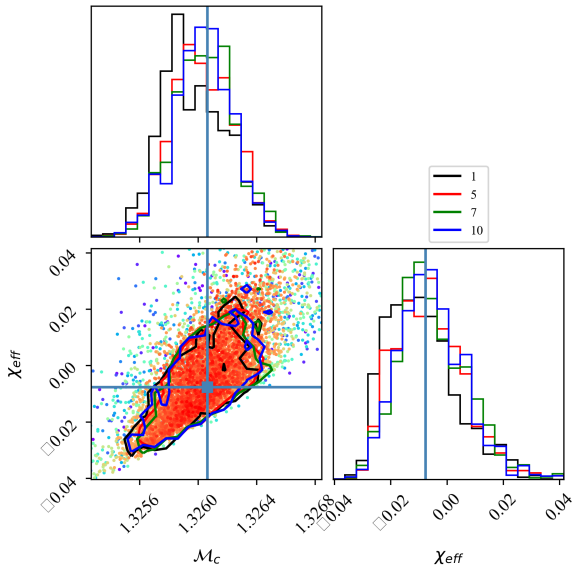

Figure 4: Demonstration of RIFT convergence over iterations. The injection is a slowly-spinning aligned-spin BNS merger denoted by the blue cross-hairs. As the analysis progresses we see the posterior converging closer to the true value from iteration 1 to 10.

siren [25].

An example of a gravitational wave signal from a compact object merger is plotted in Fig. 5. As obtaining numerical solutions to Einstein's equations describing a merger of two compact objects is computationally expensive, several analytical and semi-analytical approximations/estimates to these solutions exist. We use these waveform models/approximants to estimate parameters of a GW signal. The various waveform models that exist depending on the approximation or underlying physics being considered are [27]:

- **Numerical relativity solutions**: These are the most accurate solutions that exist till date and model all three phases of a gravitational-wave chirp- inspiral, merger and ringdown (IMR), though the inspiral cycles are just a few in number [28, 29, 30]. The merger-ringdown waveforms obtained from Numerical Relativity(NR) are used in the construction and verification of accurate waveform models used in GW data analysis.

- **Effective-one-body (EOB) approximation**: This approximation maps the dynamics of two bodies to

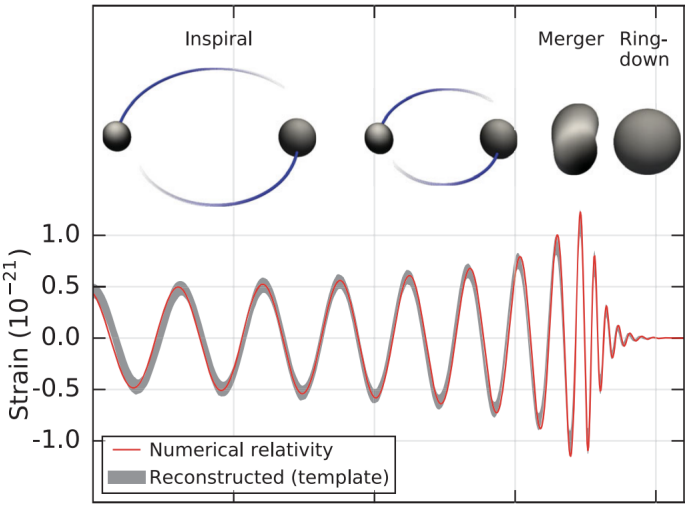

Figure 5: Estimated GW strain amplitude as a function of time from GW150914 [26] projected onto the LIGO Hanford detector (H1). The inset images show numerical relativity models of the black hole horizons as the black holes coalesce.

an effective-one-body system. These are time-domain models and includes highly accurate quadrupolar models for non-spinning and aligned-spin binaries, NR-calibrated extensions to higher-order harmonics, and extensions to precessing BBH [31, 32, 33, 34, 35]. `SEOBNRv4`, `SEOBNRv4HM` and `SEOBNRv4PHM` are some examples of models constructed using EOB approximation, HM indicating the model incorporates higher-order multipoles and P indicating a model having precession effects.

- **Phenomenological models**: This formalism models the IMR GW signal without actually providing equations of motions for the BH dynamics [36, 37]. These models are generally in frequency-domain and hence are considerably faster than EOB models. Examples approximants are `IMRPhenomD` and `IMRPhenomXAS`.

- **Numerical Relativity Surrogates**: As the name suggests, these models are interpolants of the NR simulations and are considered the most accurate merger models [38, 39, 40] but due to their restrictions in length and parameter space coverage, their usage is currently limited. `NRSur7dq4` is an example waveform model of this.

- **Analytical models**: Models like `TaylorF` (frequency-domain) [41] or `TaylorT` (time-domain)

employ post-Newtonian formalism, which is an approximation to GR valid in the slow-motion, weak-field regime. The motion of compact binaries and their GW emission in the inspiral regime is well explained by this approximation.

Waveform models that also include the effects of tides in them, used to study BNS/BHNS systems are called tidal waveform models.

1.6 Binaries with neutron star(s) components

Compact binary systems with neutron star(s) are additionally informative in providing information about the internal structure of the neutron star, hence helping in obtain better constrains on the equation of state (EOS) of the neutron star. Assuming a circular orbit of radius R for a binary of total mass M, the time the system takes to merge, also called chirp time, is roughly $\sim \frac{5M}{96}\frac{1}{\eta}(\frac{M}{R})^{-4}$, where η is the symmetric mass ratio defined as $m_1 m_2/M^2$. Hence, lower-mass systems merge on larger time scales in general. Also systems with large mass ratios between the components spend a long time in highly relativistic orbits, whereas equal-mass binaries can be expected to merge after only a few orbits in the highly relativistic regime [42]. BNS/BHNS systems will have longer inspirals compared to BBH, and these long inspirals are visible in the LIGO-Virgo frequency range. Detecting the merger or post-merger of such signals is currently hard given their frequency range limits, but future ground-based detectors such as Cosmic Explorer (CE) [43] or Einstein Telescope (ET) [44], which will have a higher maximum frequency could possibly see these signatures.

The dominant mode in a gravitational wave is the quadrupole moment or ($\ell = 2$, $m = 2$) mode, but there can also be higher-order-modes (HOMs) present ($\ell \geq 2$, $m = -\ell, ..., +\ell$), just as in sound waves. For example, GW190412 [45] which was a merger of two black holes that were highly asymmetric in masses, $30M_\odot$ and $8M_\odot$, provided the first evidence of existence of HOM in signals from binary mergers. These HOMs are sub-dominant and seen in the inspiral phase of a gravitational wave and occur at frequencies lower ($\ell = 2$, $m = 1$ mode) or higher ($\ell > 2$) than the frequency of the leading-order quadrupole mode. The inspiral phase for a BNS is longer, but weaker in amplitude, and including HOMs could help in using information from those frequencies to possibly increase the signal-to-noise ratio which in turn could reduce the statistical error bars in inferred parameter posteriors [46]. Understanding the effect of these higher

multipole moments is also important given there are correlations between masses, spins and tides which could bias the process of parameter estimation.

1.7 Waveform model systematics

As illustrated most recently by GW190521 [47, 48], GW190814 [49], GW190412 [45], and the discussion in GWTC-2 [9], these approximations disagree more than enough to produce noticeable differences, consistent with prior work [50, 51, 52]. Despite ongoing generation of new waveforms with increased accuracy [53, 54, 55, 40, 56, 57], these previous investigations suggest that waveform model systematics can remain a limiting factor in inferences about individual events [50] and populations [58, 52].

Recently, Ashton and Khan [59] described and illustrated marginalizing between a discrete set of waveform models in a fully Bayesian way. In this procedure, the waveform-marginalized posterior is the weighted average of the posteriors $p_k(\theta)$ derived from each waveform model k alone, weighted by the evidence Z_k for (and prior p_k for) each model k: $p(\theta) = [\sum_k p_k(\theta) p_k Z_k] / \sum_q p_q Z_q$. This extremely simple procedure faces one obvious limitation: analysis must be performed for every waveform model of interest. Unfortunately, as many of the most accurate time-domain waveform models incur exceptionally high evaluation costs, and as most conventional parameter estimation (PE) engines like LALInference [60] or BILBY [61] are limited by this cost, the universe of possible waveforms must often omit the most expensive and accurate waveform models. As the RIFT parameter inference engine circumvents several issues associated with waveform evaluation cost [14, 16], despite retaining the original waveform implementation (i.e., no surrogate generation), in this work we examine novel extensions of this waveform-marginalization technique which are uniquely adapted to RIFT's algorithm. Using a simple toy model, we demonstrate the pernicious effects of model systematics, then show how our technique efficiently mitigates them.

2 Binary Black Hole (BBH) gravitational waveform model systematics

2.1 Waveform models

In this work, we employ two well-studied models for non-precessing binaries, whose differences are known to be significant. We use `SEOBNRv4` [55], an effective-one-body model for quasi-circular inspiral, and `IMRPhenomD` [62, 63], a phenomenological frequency-domain inspiral-merger-ringdown model.

The effective-one-body (EOB) approach models the inspiral and spin dynamics of coalescing binaries via an ansatz for the two-body Hamiltonian [64], whose corresponding equations of motion are numerically solved in the time domain. For non-precessing binaries, outgoing gravitational radiation during the inspiral phase is generated using an ansatz for resumming the post-Newtonian expressions for outgoing radiation including non-quasicircular corrections, for the leading-order $\ell = 2$ subspace. For the merger phase of non-precessing binaries, the gravitational radiation is generated via a resummation of many quasinormal modes, with coefficients chosen to ensure smoothness. The final BH's mass and spin, as well as some parameters in the non-precessing inspiral model, are generated via calibration to numerical relativity simulations of BBH mergers.

The `IMRPhenomD` model is a part of an approach that attempts to approximate the leading-order ($\ell = 2$) gravitational wave radiation using phenomenological fits to the Fourier transform of the gravitational wave strain, computed from numerical relativity simulations, effective-one-body waveforms and post-newtonian calculation [36, 65, 53]. Also using information about the final BH state, this phenomenological frequency-domain approach matches standard approximations for the post-Newtonian gravitational wave phase to an approximate, theoretically-motivated spectrum characterizing merger and ringdown.

2.2 Fiducial synthetic sources and PP tests

We will only explore the impact of systematics over a limited fiducial population. Specifically, we consider a universe of synthetic signals for 3-detector networks, with masses drawn uniformly in m_i in the region bounded by $\mathcal{M}_c/M_\odot \in [30, 60]$ and $\eta \in [0.2, 0.25]$ and with extrinsic parameters drawn uniformly in sky

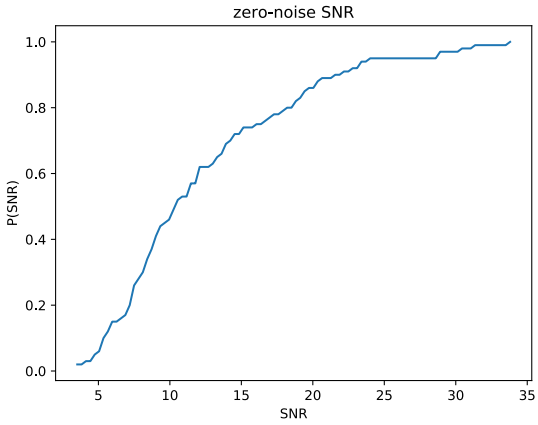

Figure 6: Cumulative SNR distribution for a synthetic population of 100 events drawn from the fiducial BBH population described in Sec. 2.2. To avoid ambiguity, this figure shows the expected SNR (i.e., the SNR evaluated using a zero-noise realization).

position and isotropically in Euler angles, with source luminosity distances drawn proportional to d_L^2 between $1.5\mathrm{Gpc}$ and $4\mathrm{Gpc}$. These bounds are expressed in terms of $\mathcal{M}_c = (m_1 m_2)^{3/5}/(m_1+m_2)^{1/5}$ and $\eta = m_1 m_2/(m_1+m_2)^2$, and encompass the detector-frame parameters of many massive binary black holes seen in GWTC-1 [8] and GWTC-2 [9]. All our sources have non-precessing spins, with each component assumed to be uniform between $[-1,1]$. For complete reproducibility, we use `SEOBNRv4` and `IMRPhenomD`, starting the signal evolution at $18\mathrm{Hz}$ but the likelihood integration at $20\mathrm{Hz}$, performing all analysis with $4096\mathrm{Hz}$ timeseries in Gaussian noise with known advanced LIGO design PSDs [66]. For each synthetic event and for each interferometer, the same noise realization is used for both waveform approximations. Ensuring convergence of the analyses, the differences between them therefore arise solely due to waveform systematics. For context, Fig. 6 shows the cumulative SNR distribution of one specific synthetic population generated from this distribution. Though a small fraction have substantial signal amplitudes, most events are near or below the level of typical detection candidates. By using a very modest-amplitude population to assess the impact of waveform systematics, we demonstrate their immediate impact on the kinds of analyses currently being performed on real observations, let alone future studies.

One way to assess the performance of parameter inference is a probability-probability plots (usually denoted PP plot) [67]. Using RIFT on each source k, with true parameters λ_k, we estimate the fraction of the

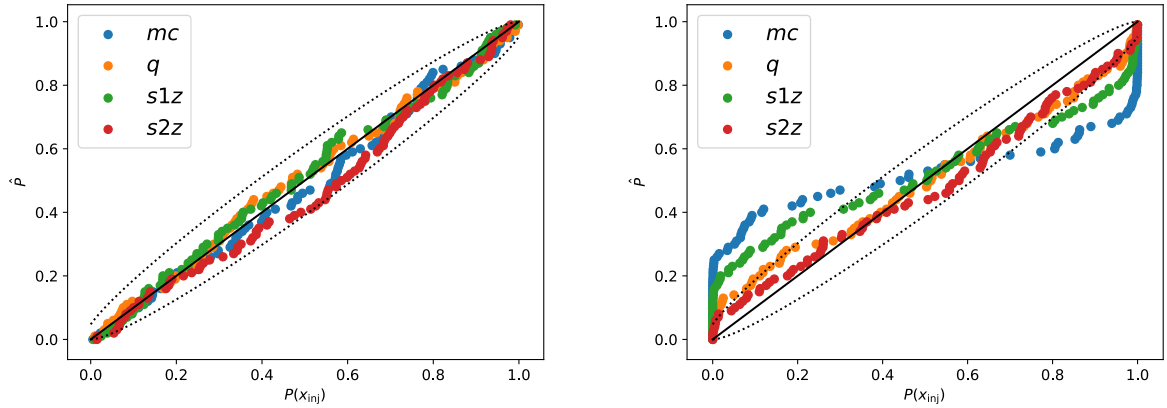

Figure 7: PP-plot of events injected with `SEOBNRv4` and recovered with `SEOBNRv4` (top panel) and `IMRPhenomD` (bottom panel) waveform. The dashed line indicates the 90% credible interval expected for a cumulative distribution drawn from 100 uniformly-distributed samples.

posterior distributions which is below the true source value $\lambda_{k,\alpha}$ [$\hat{P}_{k,\alpha}(< \lambda_{k,\alpha})$] for each intrinsic parameter α, again assuming all sources have zero spin. After reindexing the sources so $\hat{P}_{k,\alpha}(\lambda_{k,\alpha})$ increases with k for some fixed α, the top panel of Fig. 7 shows a plot of k/N versus $\hat{P}_k(\lambda_{k,\alpha})$ for all binary parameters. For the top panel, both injections and inference are performed with the same model, and the recovered probability distribution is consistent with $P(< p) = p$, as expected.

2.3 Zero noise runs to assess systematic biases

Our synthetic data consists of expected detector responses $h(t)$ superimposed on detector noise realization $n(t)$. The recovered posterior distribution's properties and in particular maximum-likelihood parameters depend on the specific noise realization used. To disentangle the deterministic effects of waveform systematics from the stochastic impact of different noise realizations, we also repeat our analyses with the "zero noise" realization: $n(t) = 0$.

2.4 Model-model mismatch

Several previous investigations (e.g., [68, 69, 70, 71, 72, 73, 52] and references therein) have phenomenologically argued that the magnitude of systematic biases are related to the model-model *mismatch*, a simple inner-product-based estimate of waveform similarity between two model predictions $h_1(\lambda)$ and $h_2(\lambda)$ at identical model parameters λ:

$$\mathcal{M}(\lambda) = 1 - \max_{t_c,\phi_c} \frac{|\left\langle h_1 | e^{i(2\pi f t_c + \phi_c)} h_2 \right\rangle|}{|h_1||h_2|} \tag{17}$$

In this expression, the inner product $\langle a|b\rangle_k \equiv \int_{-\infty}^{\infty} 2df \tilde{a}(f)^* \tilde{b}(f)/S_{h,k}(|f|)$ is implied by the kth detector's noise power spectrum $S_{h,k}(f)$, which for the purposes of waveform similarity is assumed to be the advanced LIGO instrument, H1. In practice we adopt a low-frequency cutoff $f_{\rm min}$ so all inner products are modified to

$$\langle a|b\rangle_k \equiv 2\int_{|f|>f_{\rm min}} df \frac{[\tilde{a}(f)]^* \tilde{b}(f)}{S_{h,k}(|f|)}. \tag{18}$$

Figure 8 shows the distribution of mismatches for our synthetic population, where h_1 is generated using `SEOBNRv4` and h_2 with `IMRPhenomD`. For simplicity, we regenerate all signals at zero inclination, to avoid polarization-related effects associated with the precise emission direction. For our fiducial compact binary population, the mismatches between these two models are typically below 10^{-2}, consistent with previous reports on systematic differences between these two waveforms and with their similarity to even more accurate models and simulations [55, 74, 52]

2.5 Marginalizing over waveform systematics

Suppose we have two models A and B for GW strain, and use them to interpret a particular GW source. We have prior probabilities $p(A|\lambda)$ and $p(B|\lambda)$, characterizing our relative confidence in these two models for a source with parameters λ.[1] Suppose we have produced a RIFT analysis with each model for this event, and have marginal likelihood functions $\mathcal{L}_A(\lambda)$ and $\mathcal{L}_B(\lambda)$ evaluated at a *single* point λ. We can therefore

[1] For simplicity we will assume there are no internal model hyperparameters, but the method is easily generalized to include them.

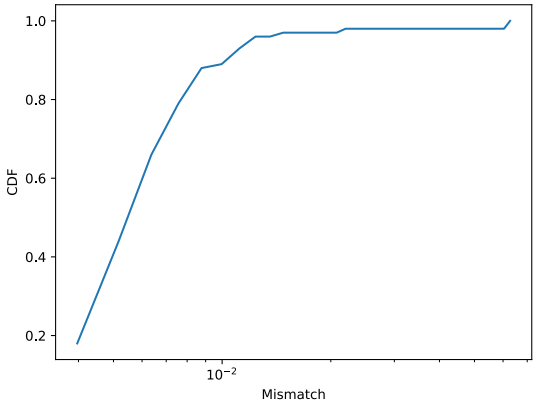

Figure 8: Cumulative mismatch distribution for all the synthetic sources in our population. We evaluate the GW strain along the z axis using `SEOBNRv4` and `IMRPhenomD` and then compute the mismatch between them. This figure shows the cumulative distribution of these mismatches, most of which are slightly less than 10^{-2}.

construct the marginal likelihood for λ by averaging over both models:

$$\mathcal{L}_{av}(\lambda) = p(A|\lambda)\mathcal{L}_A(\lambda) + p(B|\lambda)\mathcal{L}_B(\lambda) \tag{19}$$

For simplicity the calculations in this work always adopt $p(A|\lambda) = p(B|\lambda) = 1/2$. We can therefore transparently integrate multi-model inference into RIFT as follows. We assume we have a single grid of points λ_k such that both $(\lambda_k, \mathcal{L}_A(\lambda_k)$ and $(\lambda_k, \mathcal{L}_B(\lambda_k)$ can be interpolated to produce reliable likelihoods and thus posterior distributions $p_A(\lambda)$ and $p_B(\lambda)$, respectively. At each point λ_k we therefore construct $\mathcal{L}_{av}(\lambda_k)$ by the above procedure. We then interpolate to approximate $\hat{\mathcal{L}}(\lambda)$ versus the continuous parameters λ.

Operationally speaking, we construct model-averaged marginal likelihoods by the following procedure. First, we construct a fiducial grid for models A and B, for example by joining the grids used to independently analyze A and B. We use an algorithm to integrate the extrinsic likelihood (ILE), a process where each candidate GW signal is compared to a regular grid of candidate source parameters to produce an array of candidate likelihood values, to evaluate $\mathcal{L}_A(\lambda_k)$ and $\mathcal{L}_B(\lambda_k)$ on this grid [75, 14]. We construct $\mathcal{L}_{av}(\lambda_k)$ as above. We use the combinations $(\lambda_k, \mathcal{L}_{av})$ with an algorithm to construct the intrinsic posterior (CIP) from

this sampled data, to construct a model-averaged posterior distribution [14].

Our procedure bears considerable resemblance to the approach suggested by Ashton and Khan, but we have organized the calculation differently. In that approach, AK used the evidences $Z_A = \int \mathcal{L}_A p(\lambda) d\lambda$ and Z_B for the two waveform models. While we can compute both quantities with very high accuracy, we prefer to directly average between waveform models *at the same choice of intrinsic parameters* (i.e., via Eq. (19)) , to insure that marginalization over waveform models is completely decoupled from the interpolation techniques used to construct $\hat{\mathcal{L}}$ from the sampled data.

Using our fiducial BBH population, we generated 100 synthetic signals using `IMRPhenomD`, and another 100 synthetic signals with `SEOBNRv4`. For each signal, we performed parameter inference with *both* `IMRPhenomD` and `SEOBNRv4`. These inferences allow us both to assess the impact of waveform systematics in our fiducial population, and mitigate them.

2.6 Demonstrating and quantifying waveform systematics

The PP plot provides the most compelling demonstration of waveform systematics' pernicious impact. Ideally, when recovering a known model and a known population, we expect to recover the injected values as often as they occur, producing a diagonal PP plot. The top panel of Fig. 7 shows precisely what we expect, when we inject and recover with the same model (here, `SEOBNRv4`). By contrast, the bottom panel shows a PP plot generated using inference from `IMRPhenomD` on the same `SEOBNRv4` injections. The PP plot is considerably non-diagonal, reflecting frequent and substantial parameter biases in our fiducial population.

Parameter biases introduced by waveform systematics vary in magnitude and direction over the parameter space. To illustrate these offsets for the parameters $x = \mathcal{M}_c, q, \chi_{\rm eff}$, we've evaluated the parameter shift Δx between the mean inferred with `IMRPhenomD` and the mean inferred with `SEOBNRv4`, *relative* to $\sigma\rho$, which is a product of ρ (the signal-to-noise ratio, a measure of the signal amplitude) and the statistical error (as measured by the standard deviation σ of the posterior of the parameter x in question). [The combination $\sigma\rho$ is approximately independent of signal amplitude, allowing us to measure the effect of waveform systematics for a fiducial amplitude.] Figure 9 shows a vector plot of these scaled offsets $\Delta x/\rho\sigma$, as a function of two of the parameters at a time. The length of the arrow corresponds to the scaled shifts in the parameters $\mathcal{M}$, q and $\chi_{\rm eff}$, plotted against the injected parameter values. The color scale shows the remaining parameter.

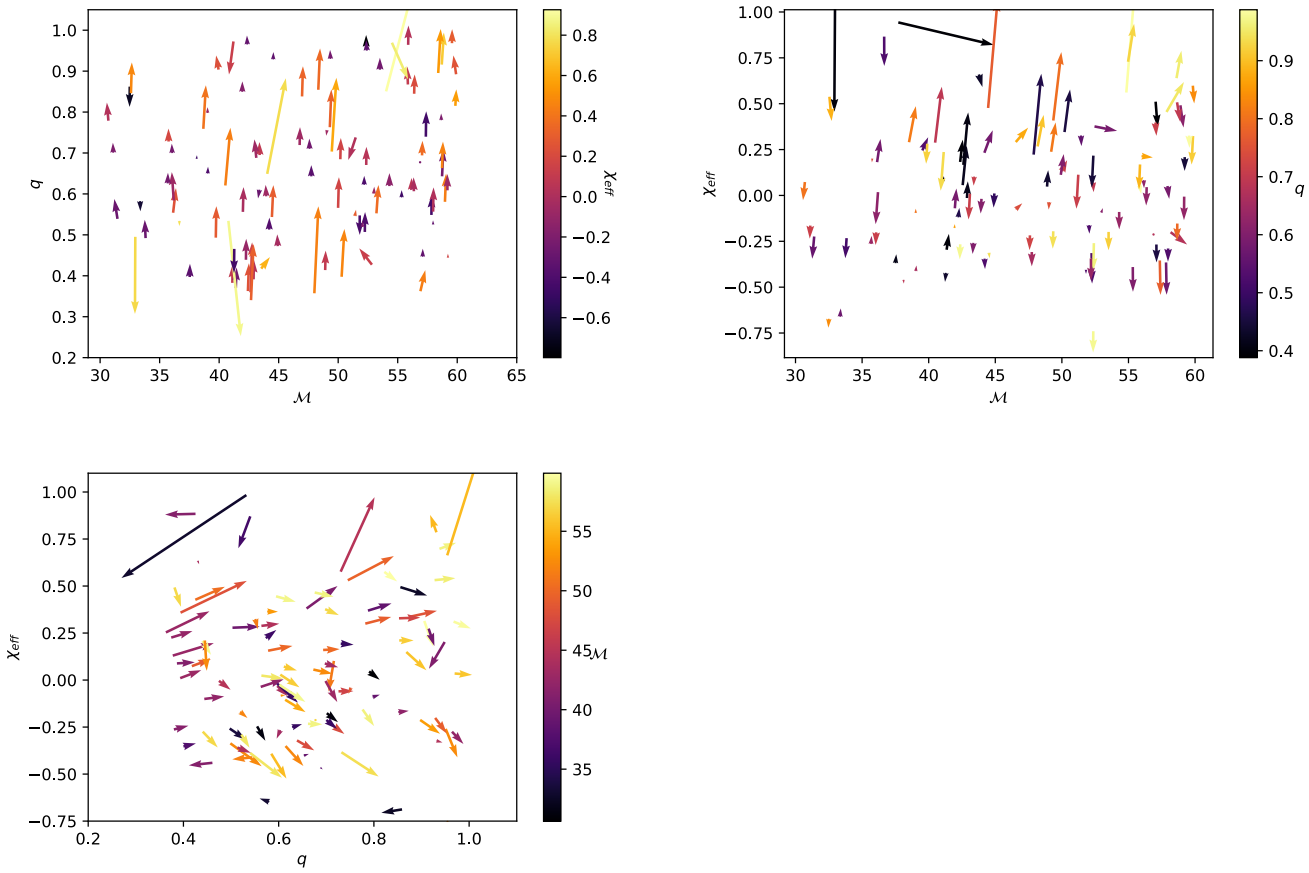

Figure 9: Vector plot showing amplitude-scaled offsets between `SEOBNRv4` and `IMRPhenomD` for parameters $\mathcal{M}$ and q (top panel), $\mathcal{M}$ and $\chi_{\rm eff}$ (middle panel) and q and $\chi_{\rm eff}$ (bottom panel) as a function of the respective parameters with color map being the value of the parameter mentioned on the color scale.

The top two panels show that shifts in $q = m_2/m_1$, $\chi_{\rm eff} = (m_1\chi_{1,z} + m_2\chi_{2,z})/(m_1 + m_2)$ are substantial. Parameter shifts for q generally increase with $\chi_{\rm eff}$. Shifts in $\chi_{\rm eff}$ are generally positive for positive $\chi_{\rm eff}$, negative for negative $\chi_{\rm eff}$, and strongly dependent on mass ratio, with more substantial shifts at either comparable mass or at very high mass ratio, respectively. In both cases, chirp mass $\mathcal{M}_c$ has modest impact, with somewhat larger shifts occurring at somewhat larger values of chirp mass. Most extreme waveform systematics seem to be associated with large mass ratio.

Analytical approximation of such shifts can help in correcting for biases introduced by the model(s)

being utilized.

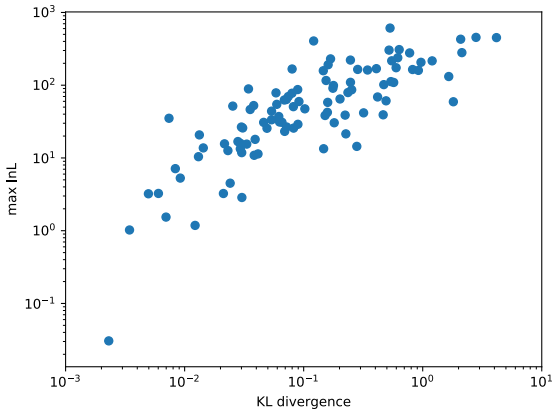

Figure 10: Figure showing KL-divergences between the two waveform models versus the log of the maximum likelihood for the combined posteriors of $\mathcal{M}$, q and $\chi_{\rm eff}$.

Relative differences in mean value only imperfectly captures the differences between the two posteriors. As a sharper diagnostic that includes parameter correlations, we use the mean and covariance of each distribution in $\mathcal{M}_c, q, \chi_{\rm eff}$ to generate a local gaussian approximation for each posterior, and then compute the Kullback–Leibler (KL) divergence between these two gaussian approximations [14]. We expect more substantial differences and thus larger KL divergence for stronger signals, whose posteriors are more sharply constrained. To corroborate our intuition, Fig. 10 shows a scatterplot, with these KL divergences on the horizontal axis and the largest value of $\ln \mathcal{L}$ on the vertical axis. As expected, for the strongest signals, differences between the two waveform models are the most pronounced.

One might expect that large parameter offsets are more likely to occur when the data favors one model or another. While conceivably true asymptotically, for our specific synthetic population, we don't find a strong correlation between the Bayes factor ($\mathcal{Z}_{\texttt{SEOBNRv4}}/\mathcal{Z}_{\texttt{IMRPhenomD}}$) and any parameter offsets. Figure 11 shows this Bayes factor (BF) plotted versus the scaled parameter offsets in $\mathcal{M}_c, q, \chi_{\rm eff}$. Large offsets can occur without the data more strongly favoring one model or the other, and vice versa.

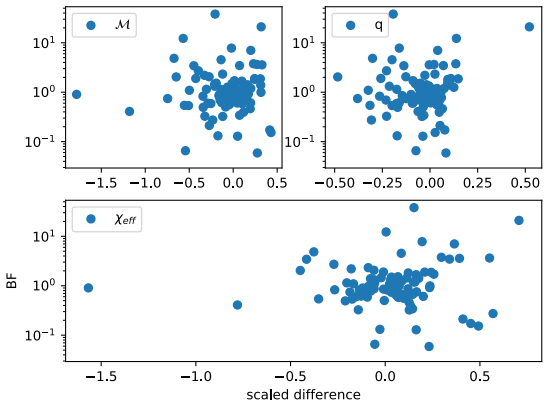

Figure 11: Figure showing Bayes factor (BF) for `SEOBNRv4` versus `IMRPhenomD` plotted against differences between the `SEOBNRv4` and `IMRPhenomD` waveforms for parameters $\mathcal{M}$, q and $\chi_{\rm eff}$.

2.7 A PP plot test for marginalizing over waveform errors

We test our model-averaged waveform procedure using a full synthetic PP plot procedure. Specifically, we use the $n_s = 100$ synthetic source parameters. For each source, we pick one waveform model A, B with probabilities $p(A), p(B)$, and use it to generate the signal. We then analyze the signal using the model-averaged procedure described above.

As a concrete example, the top panel of Fig. 12 shows our analysis of one fiducial event in our synthetic sample. The colored points show likelihood evaluations, with color scale corresponding to the marginalized likelihood evaluated with `IMRPhenomD`. The blue and black contours show the 90% credible intervals for `SEOBNRv4` and `IMRPhenomD`, respectively; the two posteriors differ substantially (i.e., the shift in mean in m_1 is of order one standard deviation), illustrating the impact of model systematics on parameter inference. The green contour shows our model-marginalized posterior. For comparison, the cross shows the injected source parameters, and the model was `IMRPhenomD`.

The bottom panel of Fig. 12 shows one PP plot corresponding to applying our model-marginalized procedure to a population where each source is randomly selected from either `IMRPhenomD` or `SEOBNRv4`. The dotted line shows a 90% frequentist interval for the largest of four random cumulative distributions. This figure shows our PP plots are consistent with the diagonal, as desired.

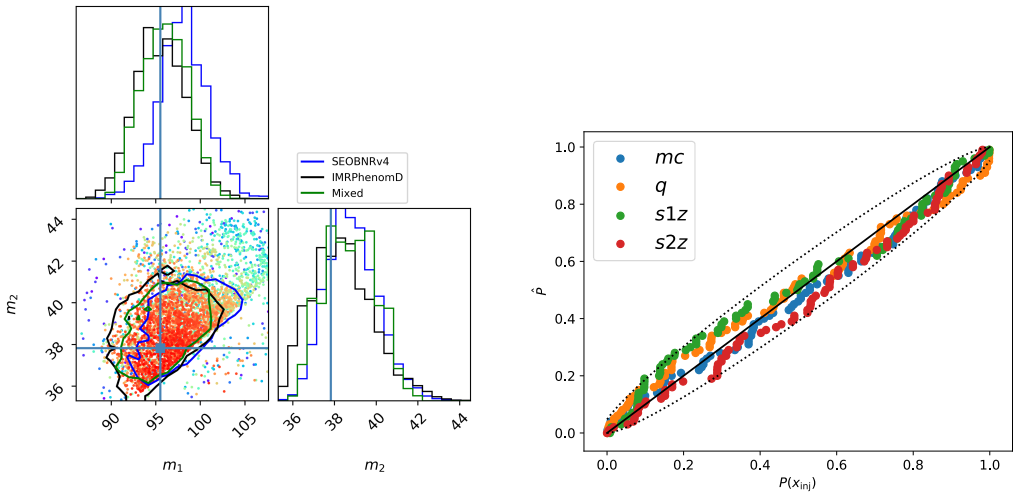

Figure 12: Top panel: example of a model averaged result. The blue and black curves show the 1D marginal distributions and 2D 90% credibles for `SEOBNRv4` and `IMRPhenomD` inferences, respectively. The green curves show the corresponding model-averaged result. Bottom panel: PP plot test for our model-marginalized procedure.

3 Binary Neutron Star (BNS) gravitational waveform model systematics

3.1 Motivation

In the LIGO-Virgo frequency band, the signal from a binary neutron-star (BNS) merger consists of the inspiral phase. The inspiral phase can be modeled well for the inclusion of higher-order-modes(HOM). Incorporation of this effect is also important as exploring a 17-dimensional parameter space with strong correlations in masses, spins and tides could lead to biased interpretation of results. Previous works done on systematics for BNS mainly include mismatch studies but are mostly done for models having only leading-order (2,2) mode [76, 77, 78]. A study done with fiducial BNS signals with HOMs argued that biases in inferring the reduced tidal parameter could be larger than the statistical 90% only for very high SNR signals $\sim$80 [79] in the LIGO-Virgo band.

In this work we simulate of a population of BNS mergers with waveform models that incorporate HOM and quantify and mitigate biases in parameter inference, if any.

3.2 Fiducial synthetic sources

The tidal waveform models used in this study are `IMRPhenomD_NRTidalv2` and `NRHybSur3dq8Tidal`. NRTidalv2 models [80] are improved versions of NRTidal [81] models, which are closed-form tidal approximants for binary neutron star coalescence and have been analytically added to selected binary black hole GW model to obtain a binary neutron star waveform, either in the time or in the frequency domain. The `NRHybSur3dq8Tidal` [82] tidal model is based on the binary black hole hybrid model `NRHybSur3dq8`, which is constructed via an interpolation of NR waveforms. It includes all modes $\ell \leq 4$, $(5, \pm 5)$ but not $(4, \pm 1)$ and (4,0) and models tidal effect up to $\lambda_{1,2} < 5000$. This model combines the accuracy of surrogate waveforms with the efficiency of PN models.

We consider a universe of 100 synthetic signals for 3-detector networks, with masses drawn uniformly in m_i in the region bounded by $\mathcal{M}_c/M_\odot \in [1.2, 1.4]$, $\eta \in [0.2, 0.25]$ and λ for each object uniformly distributed up to 1000. The extrinsic parameters are drawn uniformly in sky position and isotropically in

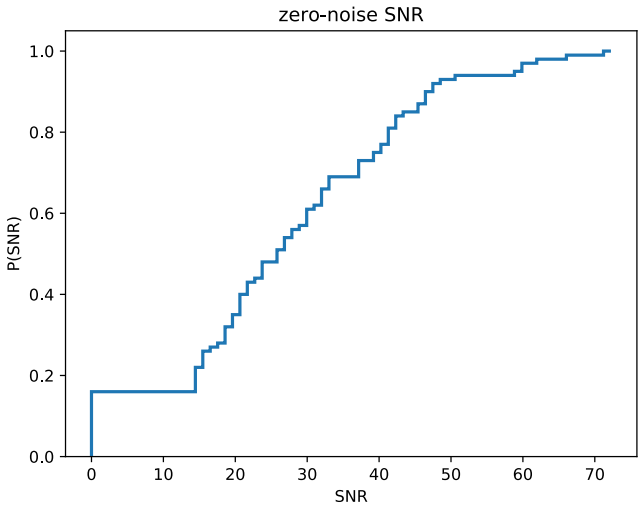

Figure 13: Cumulative SNR distribution for a synthetic population of 100 events drawn from the fiducial BNS population described in Sec. 3.2. To avoid ambiguity, this figure shows the expected SNR (i.e., the SNR evaluated using a zero-noise realization).

Euler angles, with source luminosity distances drawn proportional to d_L^2 between $30\mathrm{Mpc}$ and $80\mathrm{Mpc}$. All our sources have non-precessing spins, with each component assumed to be uniform between $[-0.05, 0.05]$. These bounds are expressed in terms of $\mathcal{M}_c = (m_1 m_2)^{3/5}/(m_1+m_2)^{1/5}$ and $\eta = m_1 m_2/(m_1+m_2)^2$, and encompass the detector-frame parameters of neutron stars observed till date. For complete reproducibility, we use `NRHybSur3dq8Tidal` and `IMRPhenomD_NRTidalv2`, starting the signal evolution and likelihood integration at $30\mathrm{Hz}$, performing all analysis with $4096\mathrm{Hz}$ time series in Gaussian noise with known advanced LIGO design PSDs [66]. The BNS signal is generated for 300 seconds but analysis performed only on 128 seconds of data. For each synthetic event and for each interferometer, the same noise realization is used for both waveform approximations. The `NRHybSur3dq8Tidal` model is utilized with two settings: a) $\ell = 5$ and b) $\ell = 2$ which includes only the dominant quadrupole mode. `IMRPhenomD_NRTidalv2` approximant has only the leading-order quadrupole mode.

Fig. 13 shows the SNR distribution (under a "zero-noise" assumption) of one specific synthetic population generated from this distribution. Comparing to GW170817's confident detection, which was a BNS merger that occurred at $40\mathrm{Mpc}$ detected by LIGO-Virgo with a SNR of 32.4, majority of the signals in this

fiducial population have SNRs below or near the typical detection criteria for a BNS merger, with some having high-enough amplitudes.

3.3 Waveform systematics

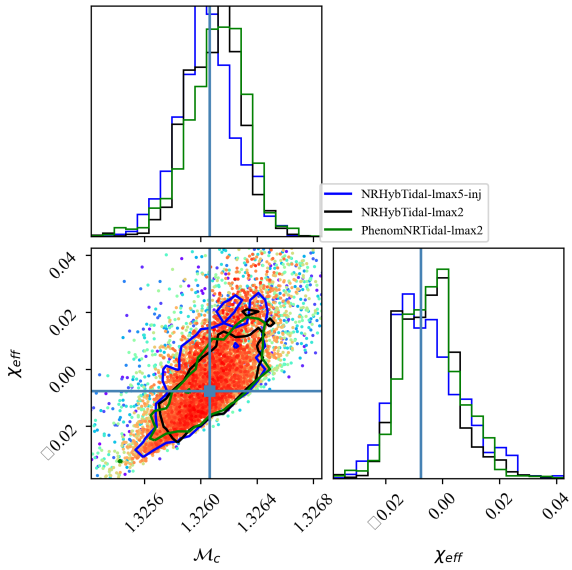

Figure 14: Corner-plot of $\mathcal{M}_c$ and χ_{eff} for a slowly-spinning aligned-spin BNS injection done with the model `NRHybSur3dq8Tidal` with HOM. Legend explains the models and settings used for parameter estimation of this signal and cross-hairs show the parameters of the injected signal.

Figure 14 shows the differences that arise in the inferred parameters of a source when higher-order-modes are included (blue) or not (black) with the same waveform approximant, `NRHybSur3dq8Tidal` of a BNS system of SNR of 24.14. There is also a clear difference in the peak of the posteriors inferred when the recovery is done with a different dominant-mode only model, `IMRPhenomD_NRTidalv2` (green). We quantify this difference in terms of relative shift (difference in the means divided by the standard deviation) in the parameters $\mathcal{M}_c$ and χ_{eff}. Between `NRHybSur3dq8Tidal` ($\ell = 5$) and `NRHybSur3dq8Tidal` ($\ell = 2$) posteriors, we see a relative shift of -0.205 in $\mathcal{M}_c$ and -0.022 in χ_{eff}. Between `NRHybSur3dq8Tidal`($\ell = 5$) and `IMRPhenomD_NRTidalv2` ($\ell = 2$) we find a shift of -

0.253 in $\mathcal{M}_c$ and -0.087 in χ_{eff}. We expect this relative shift to scale roughly as SNR, as a higher SNR signal would mean lower statistical error bars on the inferred posteriors, increasing the relative shift.

We apply the marginalization technique proposed in Sec. 2.5 to this event by considering equal weights to the two models ($p(A|\lambda) = p(B|\lambda)$= 0.5) and obtain the green posterior result shown in Fig. 15. The green posterior is obtained after performing the marginalization over `NRHybSur3dq8Tidal` ($\ell = 5$) and `IMRPhenomD_NRTidalv2` ($\ell = 2$). The green curve closely resembles the blue curve, but has a stray feature in the tail of χ_{eff} parameter as this is a preliminary result and needs to run for further iterations for that feature to be removed.

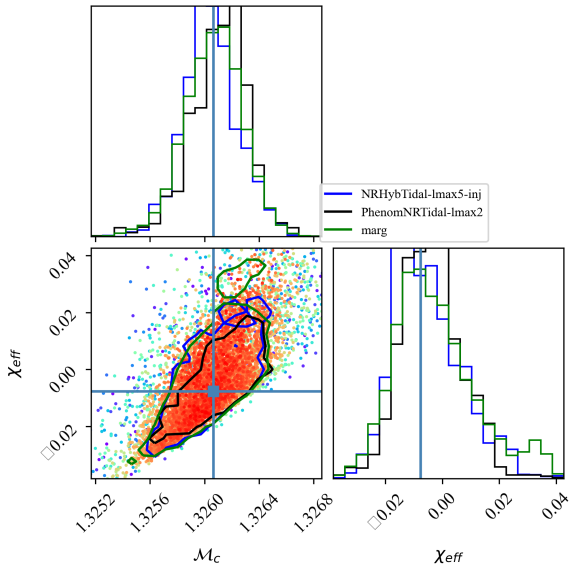

Figure 15: Corner-plot of $\mathcal{M}_c$ and χ_{eff} for a slowly-spinning aligned-spin BNS injection done with the model `NRHybSur3dq8Tidal` with HOM. Legend explains the models and settings used for parameter estimation of this signal and cross-hairs show the parameters of the injected signal. Green posterior is obtained after performing the marginalization over `NRHybSur3dq8Tidal` ($\ell = 5$) and `IMRPhenomD_NRTidalv2` ($\ell = 2$).

I am completing our analysis on the full set of synthetic events now, and will complete a paper on this subject when the results are complete. We anticipate finding results similar to our results for binary black holes: non-diagonal PP plots indicating strong systematics, and successful mitigation when we marginalize

between multiple models.

4 LIGO Scientific, Virgo, KAGRA (LVK) Collaboration work

4.1 Observation of gravitational waves from two neutron star–black hole coalescences

The Advanced LIGO and Virgo detectors observed gravitational waves from two compact binary coalescences in LIGO's and Virgo's third observing run with properties consistent with neutron star–black hole (NSBH) binaries. The two events are named GW200105_162426 and GW200115_042309, abbreviated as GW200105 and GW200115; the first was observed by LIGO Livingston and Virgo, and the second by all three LIGO–Virgo detectors. The source of GW200105 has component masses $8.9^{+1.2}_{-1.5}\,M_\odot$ and $1.9^{+0.3}_{-0.2}\,M_\odot$, whereas the source of GW200115 has component masses $5.7^{+1.8}_{-2.1}\,M_\odot$ and $1.5^{+0.7}_{-0.3}\,M_\odot$ (all measurements quoted at the 90% credible level). The probability that the secondary's mass is below the maximal mass of a neutron star is 89%–96% and 87%–98%, respectively, for GW200105 and GW200115, with the ranges arising from different astrophysical assumptions. The source luminosity distances are 280^{+110}_{-110} Mpc and 300^{+150}_{-100} Mpc, respectively. The magnitude of the primary spin of GW200105 is less than 0.23 at the 90% credible level, and its orientation is unconstrained. For GW200115, the primary spin has a negative spin projection onto the orbital angular momentum at 88% probability. We are unable to constrain the spin or tidal deformation of the secondary component for either event. We infer an NSBH merger rate density of 45^{+75}_{-33} Gpc^{-3} yr^{-1} when assuming that GW200105 and GW200115 are representative of the NSBH population, or 130^{+112}_{-69} Gpc^{-3} yr^{-1} under the assumption of a broader distribution of component masses.

Our contribution: The primary results reported in the discovery paper [3] combine samples from two BBH waveform models- `IMRPhenomXPHM` [56] and `SEOBNRv4PHM` [57] that include higher-order-modes and spin-precession effects. We contributed to the paper by inferring source properties with SEOBNRv4PHM using the parameter estimation tool RIFT. Other models which describe tides were also used to perform parameter estimation by different groups, but the primary results described in the paper are from BBH models and the explanation for this choice is provided in 4.1.1.6.

Table 1: Source properties of GW200105 and GW200115. We report the median values with 90% credible intervals. Parameter estimates are obtained using the Combined PHM samples.

	GW200105		GW200115	
	Low Spin ($\chi_2 < 0.05$)	High Spin ($\chi_2 < 0.99$)	Low Spin ($\chi_2 < 0.05$)	High Spin ($\chi_2 < 0.99$)
Primary mass $m_1/M_\odot$	$8.9^{+1.1}_{-1.3}$	$8.9^{+1.2}_{-1.5}$	$5.9^{+1.4}_{-2.1}$	$5.7^{+1.8}_{-2.1}$
Secondary mass $m_2/M_\odot$	$1.9^{+0.2}_{-0.2}$	$1.9^{+0.3}_{-0.2}$	$1.4^{+0.6}_{-0.2}$	$1.5^{+0.7}_{-0.3}$
Mass ratio q	$0.21^{+0.06}_{-0.04}$	$0.22^{+0.08}_{-0.04}$	$0.24^{+0.31}_{-0.08}$	$0.26^{+0.35}_{-0.10}$
Total mass $M/M_\odot$	$10.8^{+0.9}_{-1.0}$	$10.9^{+1.1}_{-1.2}$	$7.3^{+1.2}_{-1.5}$	$7.1^{+1.5}_{-1.4}$
Chirp mass $\mathcal{M}/M_\odot$	$3.41^{+0.08}_{-0.07}$	$3.41^{+0.08}_{-0.07}$	$2.42^{+0.05}_{-0.07}$	$2.42^{+0.05}_{-0.07}$
Detector-frame chirp mass $(1+z)\mathcal{M}/M_\odot$	$3.619^{+0.006}_{-0.006}$	$3.619^{+0.007}_{-0.008}$	$2.580^{+0.006}_{-0.007}$	$2.579^{+0.007}_{-0.007}$
Primary spin magnitude χ_1	$0.09^{+0.18}_{-0.08}$	$0.08^{+0.22}_{-0.08}$	$0.31^{+0.52}_{-0.29}$	$0.33^{+0.48}_{-0.29}$
Effective inspiral spin parameter $\chi_{\rm eff}$	$-0.01^{+0.08}_{-0.12}$	$-0.01^{+0.11}_{-0.15}$	$-0.14^{+0.17}_{-0.34}$	$-0.19^{+0.23}_{-0.35}$
Effective precession spin parameter χ_p	$0.07^{+0.15}_{-0.06}$	$0.09^{+0.14}_{-0.07}$	$0.19^{+0.28}_{-0.17}$	$0.21^{+0.30}_{-0.17}$
Luminosity distance $D_{\rm L}/{\rm Mpc}$	280^{+110}_{-110}	280^{+110}_{-110}	310^{+150}_{-110}	300^{+150}_{-100}
Source redshift z	$0.06^{+0.02}_{-0.02}$	$0.06^{+0.02}_{-0.02}$	$0.07^{+0.03}_{-0.02}$	$0.07^{+0.03}_{-0.02}$

4.1.1 Source Properties

We infer the physical properties of the two GW events using a coherent Bayesian analysis following the methodology described in Appendix B of [8]. For GW200105, data from LIGO Livingston and Virgo are analyzed, whereas for GW200115, data from both LIGO detectors and Virgo are used.

Owing to the different signal durations, we analyze 32 s of data for the higher-mass event GW200105 and 64 s of data for GW200115. All likelihood evaluations use a low-frequency cutoff of $f_{\rm low} = 20\,{\rm Hz}$, except for LIGO Livingston for GW200115, where $f_{\rm low} = 25\,{\rm Hz}$ avoids excess noise localized at low frequencies. The power spectral density used in the likelihood calculations is the median estimate calculated with BayesLine [83].

The parallel Bilby (PBILBY) inference library, together with the DYNESTY nested sampling software [61, 84, 85, 86] is the primary tool used to sample the posterior distribution of the sources' parameters and perform hypothesis testing. In addition, we use RIFT [14] for the most computationally expensive analyses and LALINFERENCE [60] for verification.

We base our main analyses of GW200105 and GW200115 on BBH waveform models that include the effects of spin-induced orbital precession and higher-order multipole GW moments, but do not include tidal

effects on the secondary. Specifically, we use two signal models: `IMRPhenomXPHM` (Phenom PHM;[56]) from the phenomenological family and `SEOBNRv4PHM` (EOBNR PHM;[57]) from the effective one-body numerical relativity family. The acronym PHM stands for Precessing Higher-order multipole Moments. Henceforth, we will use the shortened names for the waveform models.

In order to quantify the impact of neglecting tidal effects, we also analyze GW200105 and GW200115 using two NSBH waveform models that include tidal effects and assume that spins are aligned with the orbital angular momentum: `IMRPhenomNSBH` (Phenom NSBH;[87]) and `SEOBNRv4_ROM_NRTidalv2_NSBH` (EOBNR NSBH;[88]). We restrict the NSBH analyses to the region of applicability of the NSBH models, i.e. $\chi_1 < 0.5, \chi_2 < 0.05$ for Phenom NSBH and $\chi_1 < 0.9, \chi_2 < 0.05$ for EOBNR NSBH. We also perform aligned-spin BBH waveform analyses and find good agreement with the analyses using NSBH waveform models (see §4.1.1.6 below), validating the use of BBH waveform models. Specifically, we use the aligned-spin BBH models `IMRPhenomXAS` (Phenom;[89]) and `SEOBNRv4` (EOBNR;[34]), which only contain dominant quadrupole moments, and `IMRPhenomXHM` (Phenom HM;[90]) and `SEOBNRv4HM` (EOBNR HM;[91, 92]), which contain higher-order moments.

The secondary objects are probably NSs based on mass estimates. As in earlier GW analyses [6, 7], we proceed with two different priors on the secondary's spin magnitude: a *low-spin prior*, $\chi_2 \leq 0.05$, which captures the maximum spin observed in Galactic BNSs that will merge within a Hubble time [93], and a *high-spin prior*, $\chi_2 \leq 0.99$, which is agnostic about the nature of the compact object. The two priors allow us to investigate whether the astrophysically relevant subcase of low NS spin leads to differences in the parameter estimation for the binaries. All other priors are set as in previous analyses (e.g.,[9]). Throughout, we assume a standard flat ΛCDM cosmology with Hubble constant $H_0 = 67.9\ ^{-1}\,\mathrm{Mpc}^{-1}$ and matter density parameter $\Omega_\mathrm{m} = 0.3065$ [94].

For each spin prior, we run our main analyses with higher-order multipole moments and precession for both waveform families, EOBNR PHM and Phenom PHM. The EOBNR PHM model is used in combination with RIFT and the Phenom PHM model with `PBILBY`. The parameter estimation results for the individual precessing waveform models yield results in very good agreement; the median values typically differ by $1/10$ of the width of the 90% credible interval. Nevertheless, in order to alleviate potential biases due to different samplers or waveform models, we combine an equal number of samples of each into one data set

for each spin prior [95, 59, 49] and denote these as Combined PHM. The quoted parameter estimates in the following sections are the Combined PHM high-spin prior analyses. In the figures, we emphasize the high-spin prior results. The values of the most important parameters of the binaries are summarized in Table 1, and we will present details in the following sections.

4.1.1.1 Masses

Figure 16 shows the posterior distribution for the component masses of the two binaries. Defining the mass parameters such that the heavier mass is the primary object, i.e. $m_1 > m_2$, our analysis shows that GW200105 is a binary with a mass ratio of $q = m_2/m_1 = 0.22^{+0.08}_{-0.04}$, with source component masses $m_1 = 8.9^{+1.2}_{-1.5}\,M_\odot$ and $m_2 = 1.9^{+0.3}_{-0.2}\,M_\odot$. Similarly, GW200115 is a binary with a mass ratio of $q = 0.26^{+0.35}_{-0.10}$, with source component masses $m_1 = 5.7^{+1.8}_{-2.1}\,M_\odot$ and $m_2 = 1.5^{+0.7}_{-0.3}\,M_\odot$.

The primary components of GW200105 and GW200115 are identified as BHs from their mass measurements. For GW200115, we find that the probability of the primary falling in the lower mass gap ($3\,M_\odot \leq m_1 \leq 5\,M_\odot$;[99, 100]) is 30% (27%) for high-spin (low-spin) prior. For context, Fig. 16 also includes two potential NSBH candidates discovered previously; GW190814 [101] is a high-S/N event with well-measured masses that has a significantly more massive primary and a distinctly more massive secondary than either GW200105 or GW200115, and the marginal candidate GW190426_152155 [9], has (if of astrophysical origin) m_1–m_2 contours that overlap those of GW200115. The masses of GW190426_152155 are less constrained than those of GW200115 due to its smaller S/N. To highlight how the secondary masses of GW200105 and GW200115 compare to the maximum NS mass, we also show two estimates of the maximum NS mass based on an analyses of non-rotating [96] and Galactic [98] NSs.

The secondary masses are consistent with the maximum NS mass, quantified in Section 5 of the discovery paper[3].

4.1.1.2 Sky location, distance, and inclination

We localize GW200105's source to a sky area of 7200 deg^2 (90% credible region). The large sky area arises due to the absence of data from LIGO Hanford. The luminosity distance of the source is found to be $D_{\rm L} = 280^{+110}_{-110}$ Mpc. For the second event, GW200115, we localize its source to be within 600 deg^2.

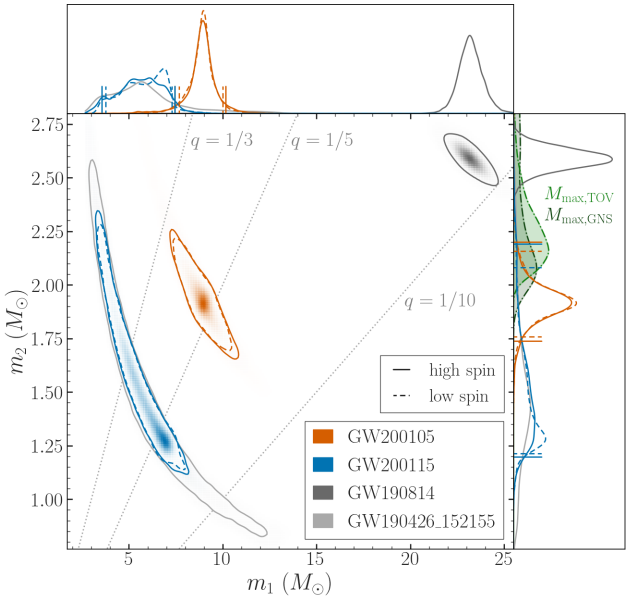

Figure 16: Component masses of GW200105 (red) and GW200115 (blue), represented by their two- and one-dimensional posterior distributions. Colored shading and solid curves indicate the high-spin prior, whereas dashed curves represent the low-spin prior. The contours in the main panel, as well as the vertical and horizontal lines in the top and right panels, respectively, indicate the 90% credible intervals. Also shown in gray are two possible NSBH events, GW190814 and the marginal candidate GW190426_152155, the latter overlapping GW200115. Lines of constant mass ratio are indicated in dashed gray. The green shaded curves in the right panel represent the one-dimensional probability densities for two estimates of the maximum NS mass, based on analyses of nonrotating NSs ($M_{\rm max,TOV}$;[96, 97]) and Galactic NSs ($M_{\rm max,GNS}$;[98]).

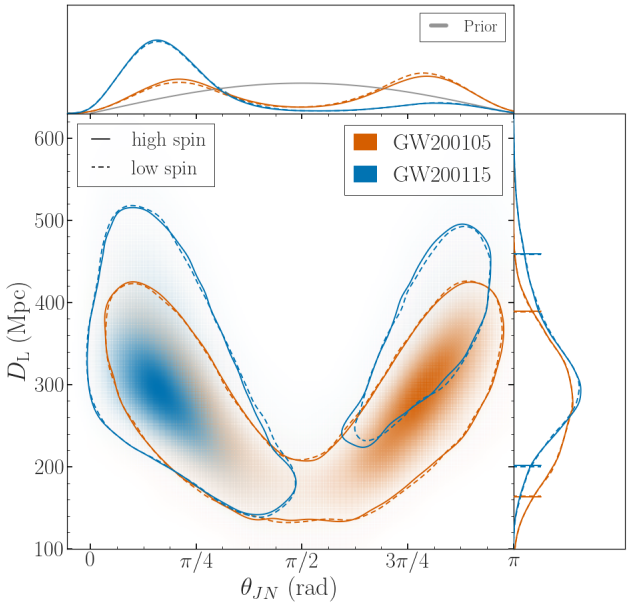

Figure 17: Two- and one-dimensional posterior distributions for distance $D_{\rm L}$ and inclinaton θ_{JN}. The solid (dashed) lines indicate the high-spin (low-spin) prior analysis, and the shading indicates the posterior probability of the high-spin prior analysis. The contours in the main panel and the horizontal lines in the right panel indicate 90% credible intervals.

It is better localized than GW200105 by an order of magnitude, since GW200115 was observed with three detectors. We find the luminosity distance of the source to be $D_{\rm L} = 300^{+150}_{-100}$ Mpc.

The luminosity distance is degenerate with the inclination angle θ_{JN} between the line of sight and the binaries' total angular momentum vector [102, 103]. Inclination $\theta_{JN} = 0$ indicates that the angular momentum vector points toward Earth. The posterior distribution of the inclination angle is bimodal and strongly correlated with luminosity distance, as shown in Fig. 17. The inclination measurement for GW200105 equally favors orbits that are either oriented toward or away from the line of sight. In contrast, GW200115 shows modest preference for an orientation $\theta_{JN} \leq \pi/2$.

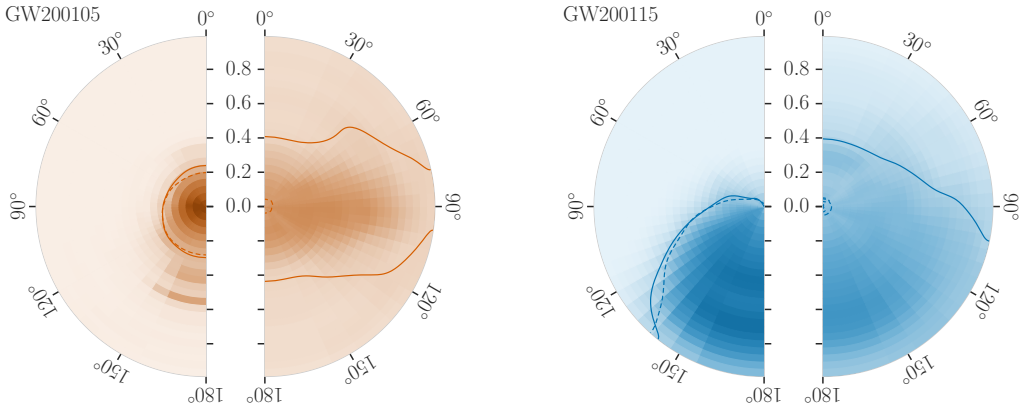

Figure 18: Two-dimensional posterior probability for the spin-tilt angle and spin magnitude for the primary objects (left hemispheres) and secondary objects (right hemispheres) for both events. Spin-tilt angles of 0° (180°) correspond to spins aligned (antialigned) with the orbital anglular momentum. The color indicates the posterior probability per pixel of the high-spin prior analysis. For comparison with the low-spin analysis, the solid (dashed) lines indicate the 90% credible regions of the high-spin (low-spin) prior analyses. The tiles are constructed linearly in spin magnitude and the cosine of the tilt angles such that each tile contains an identical prior probability. The probabilities are marginalized over the azimuthal angles.

4.1.1.3 Spins

The angular momentum vector $\vec{S}_i$ of each compact object is related to its dimensionless spin vector $\vec{\chi}_i \equiv c\vec{S}_i/(Gm_i^2)$. Its magnitude $\chi_i \equiv |\vec{\chi}_i|$ is bounded by 1. For GW200105, we infer $\chi_1 = 0.08^{+0.22}_{-0.08}$, which is consistent with zero. For GW200115, the spin magnitude is not as tightly constrained, $\chi_1 = 0.33^{+0.48}_{-0.29}$, but is also consistent with zero. The spin of the secondary for both events is unconstrained.

One of the best-constrained spin parameters is the effective inspiral spin parameter $\chi_{\rm eff}$ [104, 105, 106, 107, 108]. It encodes information about the binaries' spin components parallel to the orbital angular momentum, $\chi_{\rm eff} = \left(\frac{m_1}{M}\vec{\chi}_1 + \frac{m_2}{M}\vec{\chi}_2\right) \cdot \hat{L}$, where $\hat{L}$ is the unit vector along the orbital angular momentum.

For GW200105, $\chi_{\rm eff} = -0.01^{+0.11}_{-0.15}$ and we find the effective inspiral spin parameter to be strongly peaked about zero, with roughly equal support for being either positive or negative. For GW200115, we find modest support for negative effective inspiral spin: $\chi_{\rm eff} = -0.19^{+0.23}_{-0.35}$. Negative values of $\chi_{\rm eff}$ indicate binaries with at least one spin component negatively aligned with respect to the orbital angular momentum, i.e. $\chi_{i,z} \equiv \vec{\chi}_i \cdot \hat{L} < 0$. We find $\chi_{1,z} = -0.19^{+0.24}_{-0.50}$, and a probability of 88% that $\chi_{1,z} < 0$.

The joint posterior probability of the dimensionless spin angular momentum magnitude and tilt angle for both components of both events is shown in Fig. 18. The tilt angle with respect to the orbital angular momentum is defined as $\arccos\left(\hat{L}.\hat{\chi}_i\right)$. Deviations from uniform shading indicate a spin orientation measurement. The spin orientation of the primary of GW200105 is unconstrained, whereas the orientation of GW200115 shows support for negatively aligned primary spin.

Orbital precession is caused by a spin component in the orbital plane of a binary [109], which we parameterize using the effective precession spin parameter $0 \leq \chi_{\rm p} \leq 1$ [110]. We infer $\chi_{\rm p} = 0.09^{+0.14}_{-0.07}$ for GW200105 and $\chi_{\rm p} = 0.21^{+0.30}_{-0.17}$ for GW200115. To assess the significance of a measurement of precession, we compute a Bayes factor between a precessing and nonprecessing signal model and the precession S/N $\rho_{\rm p}$ [111, 112]. For GW200105, we find a log Bayes factor in favor of spin precession of $\log_{10}\mathcal{B} = -0.24$ and precession S/N $\rho_{\rm p} = 0.74^{+1.35}_{-0.61}$. For GW200115, $\log_{10}\mathcal{B} = -0.12$ and $\rho_{\rm p} = 0.97^{+1.57}_{-0.79}$. For both events and both diagnostics, this indicates inconclusive evidence of precession. This result is expected given the S/Ns and inferred inclination angles of the binaries [113, 114, 115].

Low values of the primary mass of GW200115 ($m_1 5\,M_\odot$) are strongly correlated with negative values of

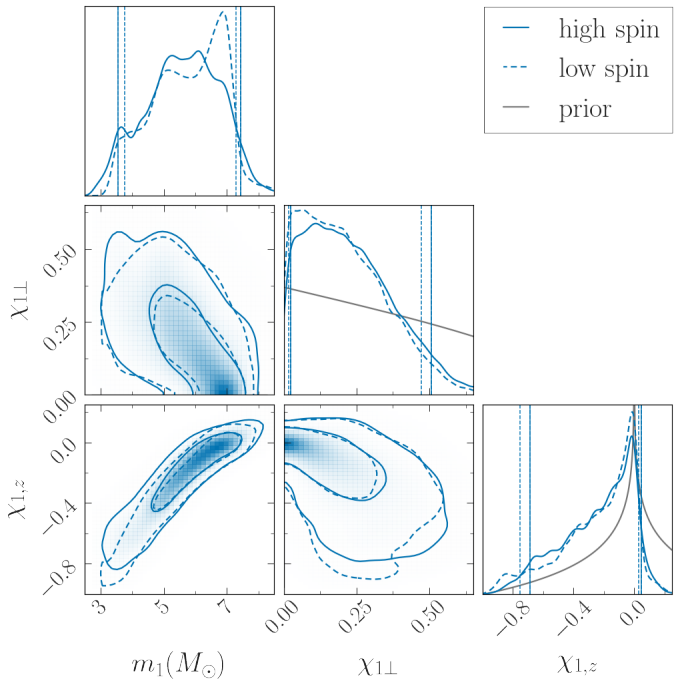

Figure 19: Properties of the primary component of GW200115. The corner plot shows the one-dimensional (diagonal) and two-dimensional (off-diagonal) marginal posterior distributions for the primary's mass and perpendicular and parallel spin components. The shading indicates the posterior probability of the high-spin prior analysis. The solid (dashed) lines indicate the 50% and 90% credible regions of the high-spin (low-spin) prior analyses. The vertical lines indicate the 90% credible intervals for the analyses with high-spin (solid lines) and low-spin (dashed lines) prior.

the primary parallel spin component $\chi_{1,z}$, as shown in Fig. 19. Figure 19 also shows the in-plane spin component $\chi_\perp$, which is peaked about zero. The lack of conclusive evidence for spin precession in GW200115 is consistent with the measurement of $\chi_\perp$. Apparent differences between the probability density of the primary spin in Fig. 18 and the posteriors of $\chi_{1\perp}$–$\chi_{1,z}$ in Fig. 19 arise from different choices in visualizing the spin orientation posteriors.

4.1.1.4 Remnant properties

Under the hypothesis of NSBH coalescence for the two events, estimates for the final mass and final spin of the remnant BH can be made using the models of [116]. We use samples obtained by combining those

from Phenom NSBH and EOB NSBH. For GW200105, the remnant mass and spin are $M_{\rm f} = 10.4^{+2.7}_{-2.0}$ and $\chi_{\rm f} = 0.43^{+0.04}_{-0.03}$, while for GW200115, $M_{\rm f} = 7.8^{+1.4}_{-1.6}$ and $\chi_{\rm f} = 0.38^{+0.04}_{-0.02}$. We do not investigate any post-merger GW signals. The S/Ns of GW200105 and GW200115 are around a factor of 3 less than that of GW170817, for which there was no evidence of GWs after the merger [117]. In the absence of tidal disruption, the postmerger signals of GW200105 and GW200115 would likely resemble a BH ringdown [118]. The GW signal associated with such ringdowns would appear well outside of LIGO's and Virgo's sensitive bandwidth given the remnant masses and spins of the systems [119].

4.1.1.5 Tests of general relativity and higher-order GW multipole moments

Results from parameterized tests of general relativity (GR;[120, 121, 122, 123, 124, 125, 126, 18]), show that GW200105 and GW200115 have too low an S/N to allow for tighter constraints than those already presented in [20]. Within their measurement uncertainties, our results do not show statistically significant evidence for deviations from the prediction of GR.

To quantify the evidence for higher-order GW multipole moments, we calculate the orthogonal optimal S/N $\rho^{\perp}_{lm}$ for the subdominant multipole moments [127, 101, 128]. We find the $(\ell, |m|) = (3, 3)$ to be the loudest subdominant multipole moment, as expected for binaries with asymmetric masses. Using the Phenom HM waveform model, we infer $\rho^{\perp}_{33} = 1.70^{+0.92}_{-1.09}$ ($1.70^{+0.94}_{-1.11}$) for GW200105 and $\rho^{\perp}_{33} = 0.91^{+0.93}_{-0.66}$ ($0.86^{+0.90}_{-0.65}$) for GW200115 with the low (high) spin prior. In Gaussian noise, the median of $\rho^{\perp}_{33}$ is approximately chi-distributed with two degrees of freedom, and values greater than 2.1 indicate significant higher-order multipole content. The measured $\rho^{\perp}_{33}$ are therefore consistent with Gaussian noise, as expected for the majority of NSBHs at these S/Ns, except for those viewed close to edge-on.

4.1.1.6 Waveform systematics

Our primary results are obtained using precessing BBH models with higher-order multipole moments, Phenom PHM and EOBNR PHM. We now justify this choice by investigating potential systematic uncertainties due to our waveform choice.

First, we investigate the agreement between independent waveform models that incorporate identical physics. Figure 20 shows the two-dimensional m_2–$\chi_{\rm eff}$ posteriors for both events obtained using a variety

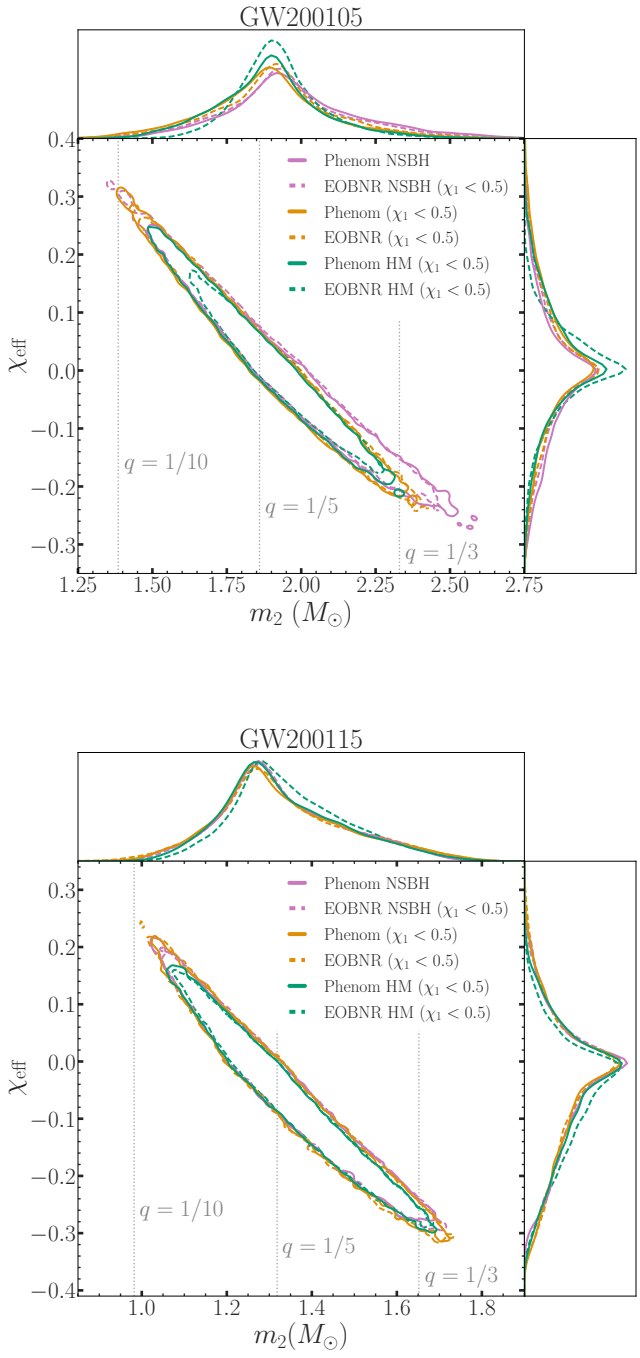

Figure 20: Comparison of two-dimensional m_2–$\chi_{\rm eff}$ posteriors for the two events reported here, using various NSBH and BBH signal models. The vertical dashed lines indicate several mass-ratio references mapped to m_2 for the median estimate of the chirp masses of GW200105 and GW200115.

of NSBH and aligned-spin BBH models. Because some NSBH models only cover $\chi_1 < 0.5$, we restrict the prior range of all models to $\chi_1 < 0.5$ for consistency.

The main panels of Fig. 20 are dominated by a correlation of the effective inspiral spin parameter $\chi_{\rm eff}$ with the secondary mass m_2 [102, 129]. Both NSBH models (Phenom NSBH and EOBNR NSBH) give consistent results with each other, as do both BBH models (Phenom and EOBNR), both with and without higher-order multipole moments, with the most notable difference being that EOBNR HM yields tighter posteriors than Phenom HM. This demonstrates that waveform models including the same physics give comparable results, but more studies are warranted to improve the understanding of the BBH waveform models in the NSBH region of parameter space. While not shown in Fig. 20, we also find good agreement between the primary precessing BBH waveform models.

Second, comparing the NSBH models with the BBH models without higher-order multipole moments (Phenom and EOBNR), the NSBH models recover similar posterior contours in the m_2–$\chi_{\rm eff}$ plane. This is expected given the asymmetric mass ratio and low S/N of these NSBH observations; see, e.g. [130] for a demonstration that higher S/Ns would be needed to see notable systematic effects. We observe differences at the extreme ends of the m_2–$\chi_{\rm eff}$ contours (i.e. at the smallest and largest values of m_2). The construction of the NSBH waveform models used here did not rely on numerical relativity results at mass ratios $q1/8$, nor did they include simulations with $\chi_{1z} < 0$ or NS masses $m_2 > 1.4 M_\odot$ [88, 87]. Therefore, some differences should be expected, especially for large m_2 in GW200105. Furthermore, for GW200105, the tails of the m_2–$\chi_{\rm eff}$ distribution for Phenom NSBH and EOB NSBH at high m_2 are also impacted by the inability of the data to constrain the tidal deformability. Hence, the posterior samples include combinations of high m_2 with large Λ_2, despite such combinations being unphysical. This effect is not apparent for GW200115 because of its smaller secondary mass. The isolated islands of probability in the extreme tails of the distributions are due to sampling noise.

Last, when adding the extra physical content of higher-order multipole moments in BBH models (through Phenom HM and EOBNR HM), the extreme ends of the m_2–$\chi_{\rm eff}$ contours are excluded, while the bulk of the distributions are consistent with the posteriors obtained with the NSBH models. In summary, these comparisons indicate that (i) waveform models including the same physics give comparable results; (ii) going from NSBH models to comparable BBH models changes the results only marginally, i.e. any effects

of tides are small; and (iii) inclusion of higher-order multipole moments changes the posterior contours more substantially than inclusion of tides. We conclude that the the inclusion of precession and higher-order multipole moments afforded by the BBH waveform models is more important than the impact of tides in the NSBH models.

Unpublished details underlying this work: The following results were done with other members of the LVK Collaboration, but were not described in the text of the paper. The statements made reflect my own conclusions and not necessarily the full opinion of the collaboration. Figure 21 (constructed from relevant public samples) shows a comparison plot of the parameters inferred for GW200115 using the two BBH-PHM waveform models-Phenom PHM and EOBNR PHM. It demonstrates possible systematics in the parameter region of this event, specially the mass and spin parameters. The EOBNR PHM model infers a more asymmetric system with smaller transverse spins (hence lesser precession in the system) and slight variation in the inferred aligned-spins as well.

Mismatches between the two models were calculated by the waveform developers in this parameter space between the two models, but were found to be generally low (0.1%-1%), concluding that at an SNR $\sim$ 12, systematics are not expected to be significant as distinguishability criterion gives mismatches of 3% would produce detectable biases. But, as shown in [2], even mismatches of the order 0.1%-1% are sufficient to show some bias, hence we are exploring the systematics for this event by performing an injection-recovery study of a fiducial event with parameters similar to that of GW200115.

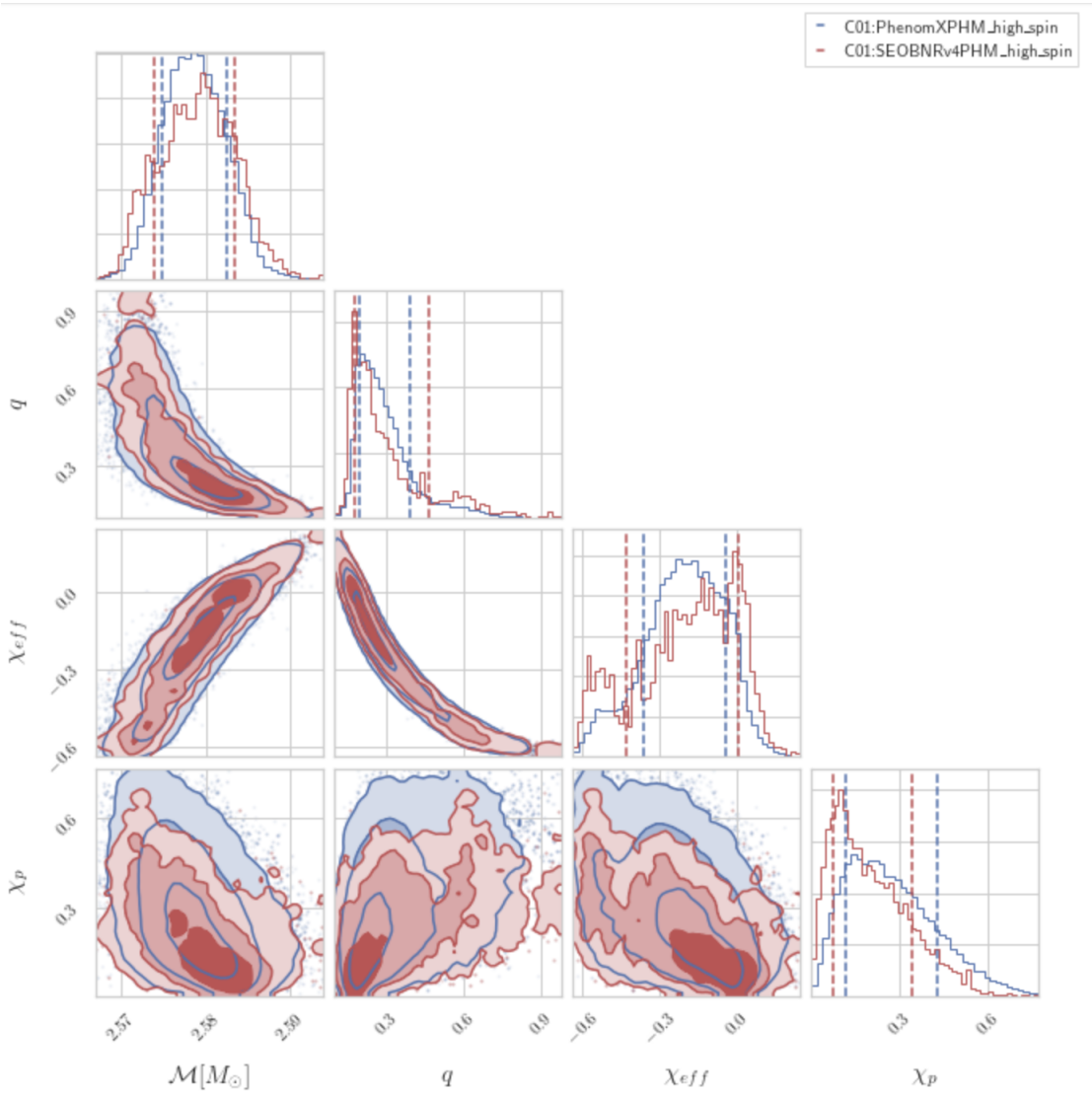

Figure 21: Corner-plot showing differences between the posterior shapes inferred using the two PHM models for GW200115 with the high-spin prior setting, blue-Phenom PHM and maroon-EOBNR PHM. $\mathcal{M}$ is the detector-frame chirp-mass and χ_p describes the amount of precession in the system. The vertical dotted lines show 90% credible intervals of the respective samples.

4.2 Search for intermediate mass black hole binaries in the third observing run of Advanced LIGO and Advanced Virgo

Intermediate-mass black holes (IMBHs) span the approximate mass range 100–$10^5\,\mathrm{M}_\odot$, between black holes (BHs) formed by stellar collapse and the supermassive BHs at the centers of galaxies. Mergers of IMBH binaries are the most energetic gravitational-wave sources accessible by the terrestrial detector network. Searches of the first two observing runs of Advanced LIGO and Advanced Virgo did not yield any significant IMBH binary signals. In the third observing run (O3), the increased network sensitivity enabled the detection of GW190521, a signal consistent with a binary merger of mass $\sim 150\,\mathrm{M}_\odot$ providing direct evidence of IMBH formation. Here we report on a dedicated search of O3 data for further IMBH binary mergers, combining both modelled (matched filter) and model independent search methods. We find some marginal candidates, but none are sufficiently significant to indicate detection of further IMBH mergers. We quantify the sensitivity of the individual search methods and of the combined search using a suite of IMBH binary signals obtained via numerical relativity, including the effects of spins misaligned with the binary orbital axis, and present the resulting upper limits on astrophysical merger rates. Our most stringent limit is for equal mass and aligned spin BH binary of total mass $200\mathrm{M}_\odot$ and effective aligned spin 0.8 at $0.056\,\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ (90% confidence), a factor of 3.5 more constraining than previous LIGO-Virgo limits. We also update the estimated rate of mergers similar to GW190521 to $0.08\,\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$.

Our contribution: A section of Appendix B of this paper [4] describes parameter estimation analysis done on event 200114_020818 which was found by the cWB offline search method with a combined FAR of $0.058\,\mathrm{yr}^{-1}$, though its terrestrial origin cannot be excluded. The network SNR with LHO-LLO network is 12.3 and the three detector network SNR is 14.5. We performed parameter estimation using RIFT with the models `SEOBNRv4PHM` and `NRSur7dq4`, both including HOMs and spin-precession effects, though NRSur7dq4 is restricted to $\frac{1}{q} > 4$ parameter space.

4.2.1 PE analysis

Here, we investigate the possibility that 200114_020818 may be described by the merger of a quasi-circular BBH system. We thus carry out parameter estimation with up-to-date waveform models including ef-

fects of precession and higher-order multipole moments. Specifically, we use three quasi-circular BBH waveform models $h(t;\theta)$: i) the numerical relativity surrogate model `NRSur7dq4` [40]; ii) the effective-one-body model `SEOBNRv4PHM` [57, 131] and iii) the phenomenological model `IMRPhenomXPHM` [56]. We perform the analysis on 8 s of data centred around 200114_020818. All analyses were performed on C01_60Hz subtracted data with a lower cutoff frequency of 10 Hz and reference frequency of 11 Hz. For the `IMRPhenomXPHM` analysis, we use the nested sampling algorithm as implemented in LALInference [60], while for `SEOBNRv4PHM` and `NRSur7dq4` analysis instead, we use the RIFT [14] analysis tool. Both algorithms are designed to compute the joint 15-dimensional posterior distribution $p(\theta|D,H)$ as well as the Bayes factor ($B_{\rm SN}$), with uniform priors on the redshifted component masses, the individual spin magnitudes, and the luminosity distance proportional to its square modulus. For the source orientation and spin vectors, we employ isotropic priors. The $\log_{10} B_{\rm SN}$ [2] is tabulated in Table 2 for all the three runs. The values of $\log_{10} B_{\rm SN}$ indicate a preference for the hypothesis H that a signal is present over the alternative of only Gaussian noise.

These results do not, though, address the possibility that excess power in one or more detectors may be due to an instrumental artifact (glitch). As a diagnostic we therefore perform a *coherence test* [132], using the `IMRPhenomXPHM` waveform model. The coherence test computes the Bayes factor for the coherent signal hypothesis against the hypothesis of an incoherent signal in the network of detectors. It can be thus interpreted loosely as an indicator of the presence of accidentally coincident noise artefacts that could mimic an astrophysical signal. The resulting $\log_{10}$ Bayes factor for coherent vs. incoherent signal 0.2, providing little to no evidence in support of the coherent signal hypothesis. Such small evidence is easily understood by looking at the $\log_{10}$ Bayes factors computed from analyses of each individual detector's data: both a Hanford-only as well as Virgo only analysis recovers a $\log_{10}$ Bayes factor for the signal vs. Gaussian noise hypothesis of 0.2. As a consequence, the posterior distributions from the Hanford- and Virgo-only analyses are largely uninformative. On the other hand, a Livingston-only analysis finds a $\log_{10}$ Bayes factor of 25. Hence, from the parameter estimation point of view, 200114_020818 is essentially a single detector event.

Returning to the results under the hypothesis H of a quasi-circular merger signal plus Gaussian noise, we summarise the resulting median and symmetric 90% credible regions for a few astrophysically relevant

[2] The uncertainties on the individual $\log_{10} B_{\rm SN}$ are ~ 1.

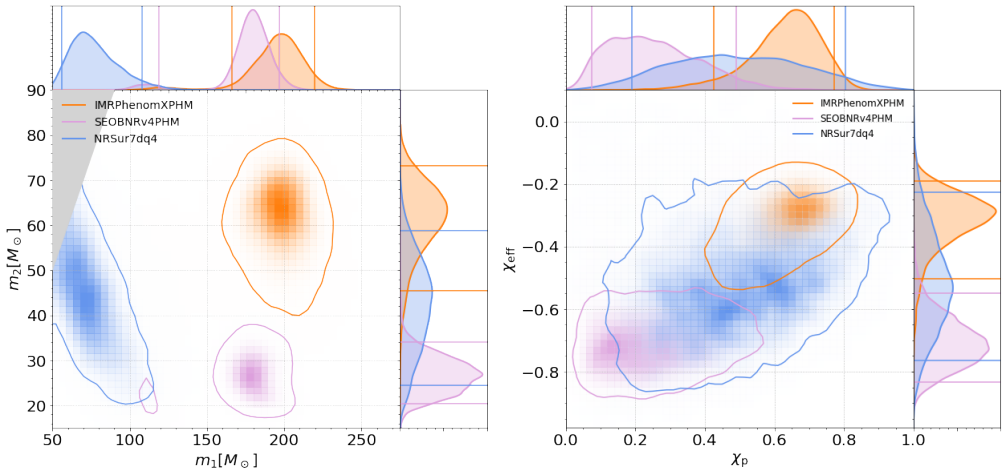

Figure 22: Posterior distributions: (Left) Source masses distribution, and (Right) the effective spin and effective in-plane spin distribution of 200114_020818 for different waveform models. The 90% credible regions are indicated by the solid contour in the joint distribution and by solid vertical and horizontal lines in the marginalized distributions.

parameters from each of the models in Table 2.

Waveform Model	m_1 ($M_\odot$)	m_2 ($M_\odot$)	$\chi_{\rm eff}$	$\chi_{\rm p}$	D_L (Mpc)	$\theta_{\rm JN}$	$\log_{10} B_{SN}$
IMRPhenomXPHM	$197.2^{+22.6}_{-31.2}$	$62.1^{+11.1}_{-16.6}$	$-0.3^{+0.1}_{-0.2}$	$0.6^{+0.1}_{-0.2}$	$593.4^{+276.8}_{-146.7}$	$0.9^{+0.8}_{-0.3}$	24.9
SEOBNRv4PHM	$179.6^{+17.1}_{-60.4}$	$26.7^{+7.3}_{-6.4}$	$-0.7^{+0.2}_{-0.1}$	$0.2^{+0.3}_{-0.2}$	$475.6^{+199.1}_{-153.0}$	$1.6^{+1.2}_{-1.1}$	26.2
NRSur7dq4	$75.0^{+32.9}_{-18.8}$	$42.5^{+16.4}_{-18.0}$	$-0.5^{+0.3}_{-0.2}$	$0.5^{+0.3}_{-0.3}$	$1797.0^{+1601.0}_{-1027.0}$	$2.0^{+0.9}_{-1.6}$	22.3

Table 2: Summary of median and 90% credible intervals of 200114_020818 for different waveform models. The columns show the waveform model used for parameter estimation, the source frame component masses m_i, effective spin parameters $\chi_{\rm eff}$ and $\chi_{\rm p}$, luminosity distance D_L, the angle between the total angular momentum and the direction of propagation of the gravitational wave signal $\theta_{\rm JN}$ and the $\log_{10}$ Bayes Factor between the signal and Gaussian Noise given the model.

Figure 22 shows the joint posterior distribution for the component masses m_1 and m_2 of the source according to each waveform model. The three models infer BH masses that are largely inconsistent. In particular, the inferred values – median and 90% credible intervals – show little overlap, see Table 2. Moreover, the result from SEOBNRv4PHM shows a hint of bimodality in the mass posterior distributions.

The posterior distributions for the spin parameters, Fig. 22, tell a similar story. If we compare the joint posterior distributions for the effective spin parameter $\chi_{\rm eff}$ along the direction of the orbital angular mo-

mentum and the in-plane effective spin parameter $\chi_{\rm p}$, we find that `IMRPhenomXPHM` and `SEOBNRv4PHM` probability distributions that are disjoint at the 90% credible level. `NRSur7dq4` instead recovers a posterior distribution that is much broader and encompasses both the posterior from `IMRPhenomXPHM` and `SEOBNRv4PHM`. With reference to Table 2, all the three results indicate a preference towards the system being precessing and with their spin vectors anti-aligned compared to the orbital angular momentum. Spin vectors anti-aligned with the orbital angular momentum have the effect of accelerating the dynamical evolution of the system towards coalescence, resulting in shorter GW signals for a given chirp mass.

In summary, a follow-up investigation of the properties of 200114_020818 interpreted as a possible quasi-circular binary merger shows considerable inconsistencies between results obtained by different waveform models. This is exemplified by the different posterior distributions for the BH masses as well as for their spins. Together with the lack of coherence among different detectors, our analysis indicates that, while we cannot exclude that 200114_020818 has an astrophysical origin, there is no consistent support for its interpretation as a quasi-circular binary merger.

5 Discussion

In this work, we demonstrated fiducial and real examples of differences in parameters inferred from binary black-hole, binary neutron-star and neutron star-black hole systems when performing parameter estimation with different waveform models. We performed simple tests which reproduce significant differences between the models `SEOBNRv4` and `IMRPhenomD`, and can be extended to other available waveforms easily using RIFT, an efficient parameter estimation engine. The probability-probability (PP) plot test, a commonly used statistical test, can be used to confirm differences between waveform models and as shown in Fig. 7, parameter estimation performed using a model different from the injected model, gives a non-diagonal PP-plot for most parameters. We calculated the magnitude and direction of the offsets introduced due to using a waveform model different to the injected model, and these differences are higher for extreme case scenarios, as expected. A linear correlation between the KL divergence computed for the two models and the log of the maximum likelihood of the injected model, shows that high-SNR signal will have larger differences in the inferred parameter from various models. Because the most informative signals exhibit the largest parameter biases, waveform systematics have the potential to strongly contaminate population inference. Most importantly, we also demonstrated a method to mitigate these waveform systematics by marginalizing over the models used for parameter estimation analyses.

We extended this analysis to a population of BNS mergers using current state-of-the-art waveform model `NRHybSur3dq8Tidal` which includes higher-order-modes(HOM) and demonstrate that inclusion/exclusion of HOM has significant effect on the parameter inference process(Fig. 14), emphasizing their importance in BNS analysis, also because this effect is dominant in the inspiral phase of a compact binary coalescence, which is what is largely observed in the LIGO-Virgo-KAGRA band for a BNS merger.

Our method requires as input some prior probabilities $p(X_k|\lambda)$ for different waveform models X_k. One way these prior probabilities could be selected is by waveform faithfulness studies between models and numerical relativity simulations. These fidelity studies inevitably suggest waveform models vary in reliability over their parameter space (e.g. [133, 134]), suggesting $p(X_k|\lambda)$ will depend nontrivially on λ. Operationally, these model priors propagate into each model's posterior inferences as if parameter inferences for model X are performed using a model-dependent prior $\propto p_{\rm prior}(\lambda)p(X_k|\lambda)$, instead of a common prior for

all models. RIFT can seamlessly perform these calculations at minimal added computational expense, while simultaneously returning results for each model derived from the conventional prior alone.

This marginalization technique can be incorporated in any Bayesian framework where you see differences when using different models or scenarios and is not designed for GW parameter estimation only. As future detectors are expected to become more sensitive, the chances of high-SNR events will be more, these systematic differences between models bring in biases, which in turn will bring in biases in inference of the population properties of the compact objects.

We further discussed the parameter estimation analysis utilized in interpretation of the first observed mergers of a neutron star-black hole and in the search for intermediate-mass black hole signals in O3 data by the LIGO, Virgo, KAGRA collaboration.

6 Conclusions

Many waveform models exist currently that describe compact binary coalescences. Even though these are derived by solving Einstein's equations, the various analytical or numerical approximation considered bring in differences and affect the parameter estimation process leading to biased interpretation of results. Averaging over the waveform models can mitigate these biases. Building on prior directly comparable work [59], we have demonstrated an efficient method to perform such model marginalization.

Other techniques have been proposed to marginalize over waveform model systematics. Notably, several groups have proposed using the error estimates provided by their model regressions (e.g., the gaussian process error) [135]. Relative to regression-based methods, our method has two notable advantages. Our method can be immediately generalized to include multiple waveform models. Critically, we plan to introduce parameter-dependent weighting of the likelihood from a waveform, since different waveforms are accurate in different regimes. No other model-marginalization technique can presently provide this level of control.

We plan to continue the study described in Sec. 3 and demonstrate the marginalization technique on a populations of BNS mergers to study the effect of inclusion of higher-order-modes for a tidal approximant. We also plan to perform a study regarding systematic differences between a model and its faster surrogate version. Future work also includes applying this technique to resolve the systematics we observed during parameter estimation analysis of the BHNS merger event- GW200115 by performing an injection-recovery study of a fiducial event with similar parameters.

References

[1] R. Abbott et al. GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. 8 2021.

[2] A. Z. Jan, A. B. Yelikar, J. Lange, and R. O'Shaughnessy. Assessing and marginalizing over compact binary coalescence waveform systematics with RIFT. *Phys. Rev. D*, 102(12):124069, 2020.

[3] R. Abbott et al. Observation of Gravitational Waves from Two Neutron Star–Black Hole Coalescences. *Astrophys. J. Lett.*, 915(1):L5, 2021.

[4] Rich Abbott et al. Search for intermediate mass black hole binaries in the third observing run of Advanced LIGO and Advanced Virgo. 5 2021.

[5] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration). Observation of Gravitational Waves from a Binary Black Hole Merger. , 116:061102–+, February 2016.

[6] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, and et al. GW170817: Observation of gravitational waves from a binary neutron star inspiral. , 119:161101, October 2017.

[7] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Gw190425: Observation of a compact binary coalescence with total mass $\sim 3.4\ m_{\odot}$. *ApJL*, 892(1):L3, March 2020.

[8] B. P. Abbott, R. Abbott, T. D. Abbott, et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *PhRvX*, 9(3):031040, July 2019.

[9] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, et al. GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. *PhRvX*, 11:021053, 2021.

[10] Jolien D. E. Creighton and Warren G. Anderson. *Gravitational-wave physics and astronomy: An introduction to theory, experiment and data analysis*. 2011.

[11] Alexander H. Nitz, Thomas Dent, Tito Dal Canton, Stephen Fairhurst, and Duncan A. Brown. Detecting binary compact-object mergers with gravitational waves: Understanding and Improving the sensitivity of the PyCBC search. *Astrophys. J.*, 849(2):118, 2017.

[12] Kipp Cannon, Sarah Caudill, Chiwai Chan, Bryce Cousins, Jolien D. E. Creighton, Becca Ewing, Heather Fong, Patrick Godwin, Chad Hanna, Shaun Hooper, Rachael Huxford, Ryan Magee, Duncan Meacher, Cody Messick, Soichiro Morisaki, Debnandini Mukherjee, Hiroaki Ohta, Alexander Pace, Stephen Privitera, Iris de Ruiter, Surabhi Sachdev, Leo Singer, Divya Singh, Ron Tapia, Leo Tsukada, Daichi Tsuna, Takuya Tsutsui, Koh Ueno, Aaron Viets, Leslie Wade, and Madeline Wade. Gstlal: A software framework for gravitational wave discovery, 2020.

[13] S. Klimenko et al. Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors. *Phys. Rev. D*, 93(4):042004, 2016.

[14] J. Lange, R. O'Shaughnessy, and M. Rizzo. Rapid and accurate parameter inference for coalescing, precessing compact binaries. *Submitted to PRD; available at arxiv:1805.10457*, 2018.

[15] J Lange. RIFT'ing the Wave: Developing and applying an algorithm to infer properties gravitational wave sources, August 2020.

[16] D. Wysocki, R. O'Shaughnessy, J. Lange, and Y.-L. L. Fang. Accelerating parameter inference with graphics processing units. , 99(8):084026, April 2019.

[17] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Tests of General Relativity with GW150914. , 116(22):221101, June 2016.

[18] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Tests of General Relativity with GW170817. , 123(1):011102, 2019.

[19] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Tests of general relativity with the binary black hole signals from the LIGO-Virgo catalog GWTC-1. , 100(10):104036, November 2019.

[20] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, and et al. Tests of general relativity with binary black holes from the second LIGO-Virgo gravitational-wave transient catalog. , 103(12):122002, 2021.

[21] D. Wysocki. Measuring the Population Properties of Merging Compact Binaries with Gravitational Wave Observations, August 2020.

[22] R. Abbott, T. D. Abbott, S. Abraham, et al. Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. , 913(1):L7, 2021.

[23] Salvatore Vitale, Ryan Lynch, Riccardo Sturani, and Philip Graff. Use of gravitational waves to probe the formation channels of compact binaries. *CQGra*, 34(3):03LT01, jan 2017.

[24] Michalis Agathos, Francesco Zappa, Sebastiano Bernuzzi, Albino Perego, Matteo Breschi, and David Radice. Inferring prompt black-hole formation in neutron star mergers from gravitational-wave data. *arXiv:1908.05442*, 2019.

[25] B. P. Abbott, R. Abbott, T. D. Abbott, et al. A gravitational-wave standard siren measurement of the Hubble constant. , 551(7678):85–88, November 2017.

[26] M. Bailes et al. Gravitational-wave physics and astronomy in the 2020s and 2030s. *Nature Rev. Phys.*, 3(5):344–366, 2021.

[27] Patricia Schmidt. Gravitational Waves From Binary Black Hole Mergers: Modeling and Observations. *Front. Astron. Space Sci.*, 7:28, 2020.

[28] Michael Boyle et al. The SXS Collaboration catalog of binary black hole simulations. *CQGra*, 36(19):195006, 2019.

[29] James Healy, Carlos O. Lousto, Jacob Lange, Richard O'Shaughnessy, Yosef Zlochower, and Manuela Campanelli. Second RIT binary black hole simulations catalog and its application to gravitational waves parameter estimation. *PhRvD*, 100(2):024021, 2019.

[30] Zachariah B. Etienne, Yuk Tung Liu, Stuart L. Shapiro, and Thomas W. Baumgarte. General relativistic simulations of black-hole–neutron-star mergers: Effects of black-hole spin. , 79:044024, Feb 2009.

[31] Thibault Damour, Alessandro Nagar, Mark Hannam, Sascha Husa, and Bernd Bruegmann. Accurate Effective-One-Body waveforms of inspiralling and coalescing black-hole binaries. , 78:044039, 2008.

[32] Enrico Barausse and Alessandra Buonanno. An Improved effective-one-body Hamiltonian for spinning black-hole binaries. , 81:084024, 2010.

[33] Michael Pürrer. Frequency domain reduced order model of aligned-spin effective-one-body waveforms with generic mass-ratios and spins. *PhRvD*, 93(6):064041, 2016.

[34] Alejandro Bohé, Lijing Shao, Andrea Taracchini, Alessandra Buonanno, Stanislav Babak, Ian W. Harry, Ian Hinder, Serguei Ossokine, Michael Pürrer, Vivien Raymond, Tony Chu, Heather Fong, Prayush Kumar, Harald P. Pfeiffer, Michael Boyle, Daniel A. Hemberger, Lawrence E. Kidder, Geoffrey Lovelace, Mark A. Scheel, and Béla Szilágyi. Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors. , 95(4):044028, 2017.

[35] Alessandro Nagar, Gunnar Riemenschneider, and Geraint Pratten. Impact of Numerical Relativity information on effective-one-body waveform models. 2017.

[36] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, J. T. Whelan, B. Brügmann, P. Diener, J. Gonzalez, M. Hannam, S. Husa, M. Koppitz, D. Pollney, L. Rezzolla, L. Santamaría, A. M. Sintes, U. Sperhake, and J. Thornburg. A phenomenological template family for black-hole coalescence waveforms. *Classical and Quantum Gravity*, 24:S689–S699, October 2007.

[37] Sebastian Khan, Katerina Chatziioannou, Mark Hannam, and Frank Ohme. Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects. *PhRvD*, 100(2):024059, 2019.

[38] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi. Numerical relativity waveform surrogate model for generically precessing binary black hole mergers. , 96(2):024058, July 2017.

[39] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, D. A. Hemberger, P. Schmidt, and R. Smith. A Surrogate model of gravitational waveforms from numerical relativity simulations of precessing binary black hole mergers. , 95(10):104023, May 2017.

[40] Vijay Varma, Scott E. Field, Mark A. Scheel, Jonathan Blackman, Davide Gerosa, Leo C. Stein, Lawrence E. Kidder, and Harald P. Pfeiffer. Surrogate models for precessing binary black hole simulations with unequal masses. *Phys. Rev. Research.*, 1:033015, 2019.

[41] Francesco Messina, Reetika Dudi, Alessandro Nagar, and Sebastiano Bernuzzi. Quasi-5.5PN TaylorF2 approximant for compact binaries: point-mass phasing and impact on the tidal polarizability inference. , 99(12):124051, 2019.

[42] B. S. Sathyaprakash and B. F. Schutz. Physics, Astrophysics and Cosmology with Gravitational Waves. *Living Rev. Rel.*, 12:2, 2009.

[43] Benjamin P Abbott et al. Exploring the Sensitivity of Next Generation Gravitational Wave Detectors. *Class. Quant. Grav.*, 34(4):044001, 2017.

[44] S. Hild et al. Sensitivity Studies for Third-Generation Gravitational Wave Observatories. *Class. Quant. Grav.*, 28:094013, 2011.

[45] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. B. Adya, and et al. GW190412: Observation of a Binary-Black-Hole Coalescence with Asymmetric Masses. , 102(4):043015, August 2020.

[46] Chris Van Den Broeck and Anand S. Sengupta. Phenomenology of amplitude-corrected post-Newtonian gravitational waveforms for compact binary inspiral. I. Signal-to-noise ratios. *Class. Quant. Grav.*, 24:155–176, 2007.

[47] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. B. Adya, and et al. GW190521: A Binary Black Hole Merger with a Total Mass of 150 M. , 125(10):101102, September 2020.

[48] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. B. Adya, and et al. Properties and astrophysical implications of the 150 Msun binary black hole merger GW190521. *arXiv e-prints*, page arXiv:2009.01190, September 2020.

[49] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. B. Adya, and et al. GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object. , 896(2):L44, June 2020.

[50] Feroz H. Shaik, Jacob Lange, Scott E. Field, Richard O'Shaughnessy, Vijay Varma, Lawrence E. Kidder, Harald P. Pfeiffer, and Daniel Wysocki. Impact of subdominant modes on the interpretation of gravitational-wave signals from heavy binary black hole systems. 11 2019.

[51] A.R. Williamson, J. Lange, R. O'Shaughnessy, J. Clark, P. Kumar, J.C. Bustillo, and J. Veitch. Inferring parameters of potentially rapidly precessing binary black holes: Systematic challenges for future graviational wave measurements. , 96:124041, December 2017.

[52] Michael Pürrer and Carl-Johan Haster. Gravitational waveform accuracy requirements for future ground-based detectors. *Physical Review Research*, 2(2):023151, May 2020.

[53] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer. Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms. , 113(15):151101, October 2014.

[54] Sebastian Khan, Katerina Chatziioannou, Mark Hannam, and Frank Ohme. Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects. , 100(2):024059, Jul 2019.

[55] A. Bohé, L. Shao, A. Taracchini, A. Buonanno, S. Babak, I. W. Harry, I. Hinder, S. Ossokine, M. Pürrer, V. Raymond, T. Chu, H. Fong, P. Kumar, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, M. A. Scheel, and B. Szilágyi. Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors. , 95(4):044028, February 2017.

[56] Geraint Pratten, Cecilio García-Quirós, Marta Colleoni, Antoni Ramos-Buades, Héctor Estellés, Maite Mateu-Lucena, Rafel Jaume, Maria Haney, David Keitel, Jonathan E. Thompson, and Sascha Husa. Let's twist again: computationally efficient models for the dominant and sub-dominant harmonic modes of precessing binary black holes. *arXiv e-prints*, page arXiv:2004.06503, April 2020.

[57] Serguei Ossokine et al. Multipolar Effective-One-Body Waveforms for Precessing Binary Black Holes: Construction and Validation. *Phys. Rev. D*, 102(4):044055, 2020.

[58] D. Wysocki, J. Lange, and R. O'Shaughnessy. Reconstructing phenomenological distributions of compact binaries via gravitational wave observations. , 100:043012, August 2019.

[59] Gregory Ashton and Sebastian Khan. Multiwaveform inference of gravitational waves. *Phys. Rev. D*, 101(6):064037, 2020.

[60] J. Veitch et al. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. *Phys. Rev. D*, 91(4):042003, 2015.

[61] Gregory Ashton, Moritz Hübner, Paul D. Lasky, Colm Talbot, Kendall Ackley, Sylvia Biscoveanu, Qi Chu, Atul Divakarla, Paul J. Easter, Boris Goncharov, Francisco Hernandez Vivanco, Jan Harms, Marcus E. Lower, Grant D. Meadors, Denyz Melchor, Ethan Payne, Matthew D. Pitkin, Jade Powell, Nikhil Sarin, Rory J. E. Smith, and Eric Thrane. BILBY: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy. , 241(2):27, April 2019.

[62] Sascha Husa, Sebastian Khan, Mark Hannam, Michael Pürrer, Frank Ohme, Xisco Jiménez Forteza, and Alejandro Bohé. Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal. , 93(4):044006, February 2016.

[63] Sebastian Khan, Sascha Husa, Mark Hannam, Frank Ohme, Michael Pürrer, Xisco Jiménez Forteza, and Alejandro Bohé. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. , 93(4):044007, February 2016.

[64] A. Taracchini, Y. Pan, A. Buonanno, E. Barausse, M. Boyle, T. Chu, G. Lovelace, H. P. Pfeiffer, and M. A. Scheel. Prototype effective-one-body model for nonprecessing spinning inspiral-merger-ringdown waveforms. , 86(2):024011, July 2012.

[65] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, E. L. Robinson, J. Seiler, and B. Krishnan. Matching post-Newtonian and numerical relativity waveforms: Systematic errors and a new phenomenological model for nonprecessing black hole binaries. , 82(6):064016–+, September 2010.

[66] LIGO Scientific Collaboration. Advanced ligo anticipated sensitivity curves. 2018.

[67] S. Cook, A. Gelman, and D. Rubin. Validation of Software for Bayesian Models Using Posterior Quantiles. *Journal of Computational and Graphical Statistics*, 15:675, 2006.

[68] L. Lindblom, B. J. Owen, and D. A. Brown. Model waveform accuracy standards for gravitational wave data analysis. , 78(12):124020–+, December 2008.

[69] Jocelyn S. Read, Charalampos Markakis, Masaru Shibata, Kōji Uryū, Jolien D. E. Creighton, and John L. Friedman. Measuring the neutron star equation of state with gravitational wave observations. , 79(12):124033, June 2009.

[70] L. Lindblom, J. G. Baker, and B. J. Owen. Improved time-domain accuracy standards for model gravitational waveforms. , 82(8):084020–+, October 2010.

[71] H.S. Cho, E. Ochsner, R. O'Shaughnessy, C. Kim, and C.H. Lee. Gravitational waves from BH-NS binaries: Phenomenological Fisher matricies and parameter estimation using higher harmonics. , 87:02400–+, January 2013.

[72] Mark Hannam, Sascha Husa, Frank Ohme, and P. Ajith. Length requirements for numerical-relativity waveforms. , 82(12):124052, December 2010.

[73] Prayush Kumar, Tony Chu, Heather Fong, Harald P. Pfeiffer, Michael Boyle, Daniel A. Hemberger, Lawrence E. Kidder, Mark A. Scheel, and Bela Szilagyi. Accuracy of binary black hole waveform models for aligned-spin binaries. , 93(10):104050, May 2016.

[74] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. , 93(4):044007, February 2016.

[75] C. Pankow, P. Brady, E. Ochsner, and R. O'Shaughnessy. Novel scheme for rapid parallel parameter estimation of gravitational waves from compact binary coalescences. , 92(2):023002, July 2015.

[76] Anuradha Samajdar and Tim Dietrich. Waveform systematics for binary neutron star gravitational wave signals: effects of the point-particle baseline and tidal descriptions. *Phys. Rev. D*, 98(12):124030, 2018.

[77] Anuradha Samajdar and Tim Dietrich. Waveform systematics for binary neutron star gravitational wave signals: Effects of spin, precession, and the observation of electromagnetic counterparts. *Phys. Rev. D*, 100(2):024046, 2019.

[78] Tim Dietrich, Tanja Hinderer, and Anuradha Samajdar. Interpreting Binary Neutron Star Mergers: Describing the Binary Neutron Star Dynamics, Modelling Gravitational Waveforms, and Analyzing Detections. *Gen. Rel. Grav.*, 53(3):27, 2021.

[79] Rossella Gamba, Matteo Breschi, Sebastiano Bernuzzi, Michalis Agathos, and Alessandro Nagar. Waveform systematics in the gravitational-wave inference of tidal parameters and equation of state from binary neutron star signals. *Phys. Rev. D*, 103(12):124015, 2021.

[80] Tim Dietrich, Anuradha Samajdar, Sebastian Khan, Nathan K. Johnson-McDaniel, Reetika Dudi, and Wolfgang Tichy. Improving the NRTidal model for binary neutron star systems. *Phys. Rev. D*, 100(4):044003, 2019.

[81] Tim Dietrich, Sebastiano Bernuzzi, and Wolfgang Tichy. Closed-form tidal approximants for binary neutron star gravitational waveforms constructed from high-resolution numerical relativity simulations. *Phys. Rev. D*, 96(12):121501, 2017.

[82] Kevin Barkett, Yanbei Chen, Mark A. Scheel, and Vijay Varma. Gravitational waveforms of binary neutron star inspirals using post-Newtonian tidal splicing. *Phys. Rev. D*, 102(2):024031, 2020.

[83] Neil J. Cornish and Tyson B. Littenberg. BayesWave: Bayesian Inference for Gravitational Wave Bursts and Instrument Glitches. *CQGra*, 32(13):135012, 2015.

[84] I. M. Romero-Shaw, C. Talbot, S. Biscoveanu, V. D'Emilio, G. Ashton, C. P. L. Berry, S. Coughlin, S. Galaudage, C. Hoy, M. Hübner, K. S. Phukon, M. Pitkin, M. Rizzo, N. Sarin, R. Smith, S. Stevenson, A. Vajpeyi, M. Arène, K. Athar, S. Banagiri, N. Bose, M. Carney, K. Chatziioannou, J. A. Clark, M. Colleoni, R. Cotesta, B. Edelman, H. Estellés, C. García-Quirós, Abhirup Ghosh, R. Green, C. J. Haster, S. Husa, D. Keitel, A. X. Kim, F. Hernandez-Vivanco, I. Magaña Hernandez, C. Karathanasis, P. D. Lasky, N. De Lillo, M. E. Lower, D. Macleod, M. Mateu-Lucena, A. Miller, M. Millhouse, S. Morisaki, S. H. Oh, S. Ossokine, E. Payne, J. Powell, G. Pratten, M. Pürrer, A. Ramos-Buades, V. Raymond, E. Thrane, J. Veitch, D. Williams, M. J. Williams, and L. Xiao. Bayesian inference for compact binary coalescences with BILBY: validation and application to the first LIGO-Virgo gravitational-wave transient catalogue. , 499(3):3295–3319, December 2020.

[85] Rory J. E. Smith, Gregory Ashton, Avi Vajpeyi, and Colm Talbot. Massively parallel Bayesian inference for transient gravitational-wave astronomy. , 498(3):4492–4502, November 2020.

[86] Joshua S Speagle. dynesty: A Dynamic Nested Sampling Package for Estimating Bayesian Posteriors and Evidences. , 493(3):3132–3158, Feb 2020.

[87] Jonathan E. Thompson, Edward Fauchon-Jones, Sebastian Khan, Elisa Nitoglia, Francesco Pannarale, Tim Dietrich, and Mark Hannam. Modeling the gravitational wave signature of neutron star black hole coalescences. , 101(12):124059, Jun 2020.

[88] Andrew Matas, Tim Dietrich, Alessandra Buonanno, Tanja Hinderer, Michael Pürrer, Francois Foucart, Michael Boyle, Matthew D. Duez, Lawrence E. Kidder, Harald P. Pfeiffer, and et al. Aligned-spin

neutron-star–black-hole waveform model based on the effective-one-body approach and numerical-relativity simulations. , 102(4):043023, 2020.

[89] Geraint Pratten, Sascha Husa, Cecilio Garcia-Quiros, Marta Colleoni, Antoni Ramos-Buades, Hector Estelles, and Rafel Jaume. Setting the cornerstone for a family of models for gravitational waves from compact binaries: The dominant harmonic for nonprecessing quasicircular black holes. , 102(6):064001, 2020.

[90] Cecilio García-Quirós, Marta Colleoni, Sascha Husa, H'ector Estell'es, Geraint Pratten, Antoni Ramos-Buades, Maite Mateu-Lucena, and Rafel Jaume. Multimode frequency-domain model for the gravitational wave signal from nonprecessing black-hole binaries. , 102(6):064002, 2020.

[91] Roberto Cotesta, Alessandra Buonanno, Alejandro Bohé, Andrea Taracchini, Ian Hinder, and Serguei Ossokine. Enriching the Symphony of Gravitational Waves from Binary Black Holes by Tuning Higher Harmonics. *PhRvD*, 98(8):084028, 2018.

[92] Roberto Cotesta, Sylvain Marsat, and Michael Pürrer. Frequency domain reduced order model of aligned-spin effective-one-body waveforms with higher-order modes. , 101(12):124040, 2020.

[93] M. Burgay, N. D'Amico, A. Possenti, R. N. Manchester, A. G. Lyne, B. C. Joshi, M. A. McLaughlin, M. Kramer, J. M. Sarkissian, F. Camilo, V. Kalogera, C. Kim, and D. R. Lorimer. An Increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system. *Nat*, 426:531–533, 2003.

[94] P. A. R. Ade, N. Aghanim, M. Arnaud, et al. Planck 2015 results. XIII. Cosmological parameters. , 594:A13, 2016.

[95] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Properties of the Binary Black Hole Merger GW150914. , 116(24):241102, June 2016.

[96] Philippe Landry, Reed Essick, and Katerina Chatziioannou. Nonparametric constraints on neutron star matter with existing and upcoming gravitational wave and pulsar observations. , 101:123007, Jun 2020.

[97] Philippe Landry, Reed Essick, and Katerina Chatziioannou. Nonparametric constraints on neutron star matter with existing and upcoming gravitational wave and pulsar observations: Weighted Monte Carlo samples for neutron star observables, April 2021.

[98] Will M. Farr and Katerina Chatziioannou. A Population-Informed Mass Estimate for Pulsar J0740+6620. *Research Notes of the American Astronomical Society*, 4(5):65, May 2020.

[99] C. D. Bailyn, R. K. Jain, P. Coppi, and J. A. Orosz. The Mass Distribution of Stellar Black Holes. , 499:367–374, May 1998.

[100] Jerome A. Orosz, Raj K. Jain, Charles D. Bailyn, Jeffrey E. McClintock, and Ronald A. Remillard. Orbital Parameters for the Soft X-Ray Transient 4U 1543-47: Evidence for a Black Hole. , 499(1):375–384, May 1998.

[101] R. Abbott, T. D. Abbott, S. Abraham, et al. GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object. , 896(2):L44, June 2020.

[102] Curt Cutler and Eanna E. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form? , 49:2658–2697, 1994.

[103] Samaya Nissanke, Daniel E. Holz, Scott A. Hughes, Neal Dalal, and Jonathan L. Sievers. Exploring Short Gamma-ray Bursts as Gravitational-wave Standard Sirens. , 725(1):496–514, Dec 2010.

[104] Thibault Damour. Coalescence of two spinning black holes: an effective one-body approach. *Phys. Rev.*, D64:124013, 2001.

[105] Etienne Racine. Analysis of spin precession in binary black hole systems including quadrupole-monopole interaction. , 78:044021, 2008.

[106] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, E. L. Robinson, J. Seiler, and B. Krishnan. Matching post-newtonian and numerical relativity waveforms: Systematic errors and a new phenomenological model for nonprecessing black hole binaries. , 82:064016, Sep 2010.

[107] P. Ajith, N. Fotopoulos, S. Privitera, A. Neunzert, and A. J. Weinstein. Effectual template bank for the detection of gravitational waves from inspiralling compact binaries with generic spins. *PhRvD*, 89(8):084041, 2014.

[108] Salvatore Vitale, Ryan Lynch, Vivien Raymond, Riccardo Sturani, John Veitch, and Philp Graff. Parameter estimation for heavy binary-black holes with networks of second-generation gravitational-wave detectors. *PhRvD*, 95(6):064053, 2017.

[109] Theocharis A. Apostolatos, Curt Cutler, Gerald J. Sussman, and Kip S. Thorne. Spin-induced orbital precession and its modulation of the gravitational waveforms from merging binaries. , 49:6274–6297, Jun 1994.

[110] Patricia Schmidt, Frank Ohme, and Mark Hannam. Towards models of gravitational waveforms from generic binaries: II. Modelling precession effects with a single effective precession parameter. , 91(2):024043, Jan 2015.

[111] Stephen Fairhurst, Rhys Green, Mark Hannam, and Charlie Hoy. When will we observe binary black holes precessing? , 102(4):041302, 2020.

[112] Stephen Fairhurst, Rhys Green, Charlie Hoy, Mark Hannam, and Alistair Muir. Two-harmonic approximation for gravitational waveforms from precessing binaries. , 102(2):024055, 2020.

[113] Salvatore Vitale, Ryan Lynch, John Veitch, Vivien Raymond, and Riccardo Sturani. Measuring the spin of black holes in binary systems using gravitational waves. *PhRvL*, 112(25):251101, 2014.

[114] Rhys Green, Charlie Hoy, Stephen Fairhurst, Mark Hannam, Francesco Pannarale, and Cory Thomas. Identifying when Precession can be Measured in Gravitational Waveforms. , 103(12):124023, 2021.

[115] Geraint Pratten, Patricia Schmidt, Riccardo Buscicchio, and Lucy M. Thomas. Measuring precession in asymmetric compact binaries. *Physical Review Research*, 2(4):043096, 2020.

[116] Francesco Zappa, Sebastiano Bernuzzi, Francesco Pannarale, Michela Mapelli, and Nicola Giacobbo. Black-Hole Remnants from Black-Hole–Neutron-Star Mergers. , 123(4):041102, 2019.

[117] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Search for Post-merger Gravitational Waves from the Remnant of the Binary Neutron Star Merger GW170817. , 851(1):L16, 2017.

[118] Francois Foucart, Luisa Buchman, Matthew D. Duez, Michael Grudich, Lawrence E. Kidder, Ilana MacDonald, Abdul Mroue, Harald P. Pfeiffer, Mark A. Scheel, and Bela Szilagyi. First direct comparison of nondisrupting neutron star-black hole and binary black hole merger simulations. , 88(6):064017, September 2013.

[119] Nikhil Sarin and Paul D. Lasky. The evolution of binary neutron star post-merger remnants: a review. *General Relativity and Gravitation*, 53(6):59, June 2021.

[120] Luc Blanchet and B.S. Sathyaprakash. Detecting the tail effect in gravitational wave experiments. , 74:1067–1070, 1995.

[121] Nicolás Yunes and Frans Pretorius. Fundamental theoretical bias in gravitational wave astrophysics and the parametrized post-einsteinian framework. *PhRvD*, 80:122003, Dec 2009.

[122] Chandra Kant Mishra, K. G. Arun, Bala R. Iyer, and B. S. Sathyaprakash. Parametrized tests of post-newtonian theory using advanced ligo and einstein telescope. *PhRvD*, 82:064010, Sep 2010.

[123] T. G. F. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos, J. Veitch, K. Grover, T. Sidery, R. Sturani, and A. Vecchio. Towards a generic test of the strong field dynamics of general relativity using compact binary coalescence. *PhRvD*, 85:082003, Apr 2012.

[124] T. G. F. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos, J. Veitch, K. Grover, T. Sidery, R. Sturani, and A. Vecchio. Towards a generic test of the strong field dynamics of general relativity using compact binary coalescence: Further investigations. *J. Phys. Conf. Ser.*, 363:012028, 2012.

[125] Michalis Agathos, Walter Del Pozzo, Tjonnie G. F. Li, Chris Van Den Broeck, John Veitch, and Salvatore Vitale. TIGER: A data analysis pipeline for testing the strong-field dynamics of general relativity with gravitational wave signals from coalescing compact binaries. , 89(8):082001, 2014.

[126] Jeroen Meidam, Ka Wa Tsang, Janna Goldstein, Michalis Agathos, Archisman Ghosh, Carl-Johan Haster, Vivien Raymond, Anuradha Samajdar, Patricia Schmidt, Rory Smith, Kent Blackburn, Walter

Del Pozzo, Scott E. Field, Tjonnie Li, Michael Pürrer, Chris Van Den Broeck, John Veitch, and Salvatore Vitale. Parametrized tests of the strong-field dynamics of general relativity using gravitational wave signals from coalescing binary black holes: Fast likelihood calculations and sensitivity of the method. , 97(4):044033, 2018.

[127] R. Abbott, T. D. Abbott, S. Abraham, et al. GW190412: Observation of a Binary-Black-Hole Coalescence with Asymmetric Masses. , 102(4):043015, 2020.

[128] Cameron Mills and Stephen Fairhurst. Measuring gravitational-wave higher-order multipoles. , 103(2):024042, 2021.

[129] Ken K. Y. Ng, Salvatore Vitale, Aaron Zimmerman, Katerina Chatziioannou, Davide Gerosa, and Carl-Johan Haster. Gravitational-wave astrophysics with effective-spin measurements: asymmetries and selection biases. *PhRvD*, 98(8):083007, 2018.

[130] Yiwen Huang, Carl-Johan Haster, Salvatore Vitale, Vijay Varma, Francois Foucart, and Sylvia Biscoveanu. Statistical and systematic uncertainties in extracting the source properties of neutron star-black hole binaries with gravitational waves. , 103(8):083001, April 2021.

[131] Stanislav Babak, Andrea Taracchini, and Alessandra Buonanno. Validating the effective-one-body model of spinning, precessing binary black holes against numerical relativity. *Phys. Rev. D*, 95(2):024010, 2017.

[132] J. Veitch and A. Vecchio. Bayesian approach to the follow-up of candidate gravitational wave signals. *Phys. Rev. D*, 78:022001, Jul 2008.

[133] Prayush Kumar, Ilana MacDonald, Duncan A. Brown, Harald P. Pfeiffer, Kipp Cannon, Michael Boyle, Lawrence E. Kidder, Abdul H. Mroué, Mark A. Scheel, Béla Szilágyi, and Anıl Zenginoğlu. Template banks for binary black hole searches with numerical relativity waveforms. , 89(4):042002, February 2014.

[134] Prayush Kumar, Kevin Barkett, Swetha Bhagwat, Nousha Afshari, Duncan A. Brown, Geoffrey Lovelace, Mark A. Scheel, and Béla Szilágyi. Accuracy and precision of gravitational-wave mod-

els of inspiraling neutron star-black hole binaries with spin: Comparison with matter-free numerical relativity in the low-frequency regime. , 92(10):102001, November 2015.

[135] Alvin J. K. Chua, Natalia Korsakova, Christopher J. Moore, Jonathan R. Gair, and Stanislav Babak. Gaussian processes for the interpolation and marginalization of waveform error in extreme-mass-ratio-inspiral parameter estimation. , 101(4):044027, February 2020.